Instituto de Agrobiotecnología. IDAB, CSIC - UPNA - Gobierno de Navarra
Departamento de Producción Agraria - Universidad Pública de Navarra

DOCTORAL THESIS

Functional analysis of the RNA chaperone CspA in *Staphylococcus aureus*

Carlos Caballero
Pamplona, 2018

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

TESIS DOCTORAL

Functional analysis of the RNA chaperone CspA in *Staphylococcus aureus*

Memoria Presentada por

CARLOS CABALLERO SÁNCHEZ

Para optar al grado de Doctor por la Universidad Pública de Navarra

Director:

Dr. ALEJANDRO TOLEDO ARANA

Pamplona, 2018

Dr. ALEJANDRO TOLEDO-ARANA, Científico Titular del Consejo Superior de Investigaciones Científicas (CSIC), responsable del Grupo de Regulación Génica Bacteriana del Instituto de Agrobiotecnología (IdAB, CSIC-UPNA-GN)

INFORMA:

Que la presente memoria de Tesis Doctoral titulada "**Functional analysis of the RNA chaperone CspA in *Staphylococcus aureus***" que ha sido elaborada por Don **CARLOS CABALLERO SÁNCHEZ**, ha sido realizada bajo su dirección y que cumple las condiciones exigidas por la legislación vigente para optar al grado de Doctor.

Y para que así conste, firma la presente en Pamplona, a 8 de Febrero de 2018

Fdo. Alejandro Toledo Arana

AGRADECIMIENTOS

AGRADECIMIENTOS

Toda etapa llega a su fin y cuando esto sucede solo nos queda echar la vista atrás y dar las gracias por todo lo vivido. A lo largo de este proceso me he sentido acompañado por muchas personas que merecen ser nombradas ya que, de un modo u otro, han colaborado en el desarrollo de este proyecto. Por ello, quiero manifestar mi más profundo agradecimiento a:

La Universidad Pública de Navarra, el Consejo Superior de Investigaciones Científicas y en especial al Instituto de Agrobiotecnología.

Alejandro Toledo-Arana, que además de dirigir este trabajo ha sido mi mentor tanto en mi formación académica como personal, por su implicación incondicional, profesionalidad, ofrecerme multitud de oportunidades, confiar en mí, su profundo sentido del respeto y su amistad. Sin él este trabajo no hubiese sido posible.

Todo el grupo de Regula Genes por su dedicación y esfuerzo, así como por regalarme su lado más humano: Pilar, Naiara, Aran Morens, Laurène, Leticia, Laura y Sergio.

Iñigo Lasa, por abrirme las puertas de su laboratorio y seguir de cerca mis progresos de forma crítica y constructiva. Del mismo modo, no quiero olvidarme de todo su equipo de biofilms, con los que compartí mis primeros años de formación: Cristina, Jaione, Bego, Maite E., Pedro, Carmen, Bea y Lorena, la biofilm en adopción. De la hermana Saioa, el maestro Igor, Sonia grumaquer y la capitana Maite Villanueva, que me guió y soportó durante mis primeros pasos.

Victor Segura por prestar su ayuda y conocimientos bioinformáticos.

Todo el grupo de Fernando Corrales, por su servicio en los análisis de proteómica.

Pascale Cossart y a todo su equipo, en especial a Javier, por su hospitalidad y cortesía durante toda mi estancia.

Pascale Romby, líder del grupo de Arquitectura y reactividad del RNA en el IBMC de Estrasburgo, por su calidez y acogida. Asimismo, a los miembros de su grupo, Delphine, Arnaud, Melodie, Pierre, Anne-Catherine, Stefano, Isabelle. También a mis colegas Alessandra y Zong-Fu W., por los buenos momentos juntos.

Mari Jose, Idoia, Marta, Maite, Maria, Alfonso, Ander, Inés y Sarah por el valor incalculable de su servicio. A Fernando A y Sara O., por sus labores administrativas, y a Victor por su entrega.

Mis amigas Argi, Leti, Anaca y a mi querida Tana, con la que tantos momentos de calidad compartí y que tanto me enseñó de la vida.

Mis hermanos "*masones*", que siempre estuvieron preocupados de que todo saliese bien, han sido y serán siempre un apoyo fundamental en mi vida. También, a toda mi familia de "*Paul Segers 5"* por mantenerse unida.

Mis hermanos, a mi madre y a mi padre. A estos últimos de forma especial porque sus sacrificios me han ofrecido unas condiciones en las que tengo la oportunidad de elegir mi futuro. Particularmente, a mi padre le dedico unas líneas porque me ha sabido orientar y aconsejar en mi proceso formativo.

Raquel, por tanto amor recibido, sobre todo en los momentos de mayor desaliento.

Este trabajo ha sido realizado gracias a la financiación recibida a través de los siguientes proyectos de investigación:

2011- “Regulación post-transcripcional mediada por las regiones 3' no traducidas del RNA mensajero en bacterias”. Ministerio de Ciencia e Innovación (BFU2011-23222)

2015- “Regulación post-transcripcional mediada por las regiones 3' no traducidas del RNA mensajero en bacterias”. Ministerio de Economía y Competitividad (BFU2011-56698-P)

2015- “Regulación post-transcripcional mediada por RNAs y proteínas de unión a RNA en bacterias patógenas”. Científico Titular del Consejo Superior de Investigaciones Científicas (PIE-201540I013)

2015- “High-throughput *in vivo* studies on post-transcriptional regulatory mechanisms mediated by bacterial 3’-UTRs”. European Research Council Consolidator Grant: ERC-CoG-2014-646869

CONTENTS

SUMMARY

SUMMARY

RNA-binding proteins (RBPs) are essential elements to fine-tune gene expression at the post-transcriptional level. They are numerous and, in a general manner, can be grouped in ribosomal proteins, ribonucleases and RNA chaperones, which comprise diverse protein domains. Cold shock proteins (CSPs) are a group of RNA chaperones belonging to the cold-shock domain (CSD) family that is present in every kingdom of life. CSPs have been extensively studied but the specific RNA targets and the biological functions remain elusive for most of them. The name CSP comes from the first identified protein, whose expression was inducible by cold. However, most of the members of this family are not cold-inducible. In addition, one genome may encode several CSPs that are suspected to be functionally redundant due to their high sequence identity.

In this Thesis, we have explored the regulon, specificity, biological role and regulatory mechanisms of the RNA chaperone CspA using *Staphylococcus aureus* as a model. *S. aureus* is a notorious pathogen that produces a wide range of diseases and, with the emergence of multidrug-resistant strains, has raised awareness towards generating new antimicrobial strategies. Its genome encodes three CSP genes with a sequence identity higher than 70% between them, suggesting a possible functional overlap. To address this matter, we selected a *cspA* mutant strain, which presented reduced staphyloxanthin (STX) production, as an *in vivo* functional reporter for complementation studies. If CSP paralogs were functionally redundant, STX levels would be restored by expressing any of *S. aureus* CSPs. Expression of CSPs under the control of a

heterologous promoter showed that only CspA restored STX production. However, Western blots revealed different CSPs levels, indicating the existence of differential post-transcriptional regulatory processes, which possibly prevented the expression of appropriate CSP amounts to complement STX production. In an attempt to avoid such post-transcriptional regulation bias, we genetically modified the *cspA* mRNA to encode CspB or CspC. Thanks to these mutated genes, CspB and CspC proteins were produced in comparable levels to those of CspA. Nevertheless, they were still unable to restore the wild type level of STX, suggesting individual functionalities for each of *S. aureus* CSP. To define the protein motifs responsible for CspA target specificity, we created CSP chimeric constructs by interchanging CspB and CspC motifs with their corresponding ones in CspA. We found that the amino acid differences in the carboxi-half of the protein, away from the RNA-binding motifs, were responsible for functional singularity.

Seeing that *S. aureus* CSPs were not interchangeable, we focused on deciphering the specific biological role of CspA by elucidating its regulon. For this purpose, we combined genome-wide proteomics with *in vivo* RIP-CHIP and found that CspA behaved as a global regulator of gene expression, modulating genes required for carbohydrate and ribonucleotide metabolism, stress response and virulence. Consequently, deletion of *cspA* deregulated biofilm formation and impeded bacterial survival in the presence of oxidative agents. This highlighted the importance of CspA for *S. aureus* to cope more efficiently with environmental stresses.

We noticed that CspA-binding could either promote or reduce protein expression. The latter was an unexpected trait since CSPs are widely regarded as enhancers of translation by unfolding RNA structures that impair ribosome progression on the mRNAs. Nevertheless, we observed that the repression exerted by CspA could occur by at least two different mechanisms. In the first case, CspA bound the 5'UTR of the *cspC* mRNA and decreased its translation without changing the mRNA levels. In the second case, CspA repressed its own expression by interfering with endoribonuclease III (RNase III) activity. RNase III is required for promoting CspA translation through processing of an RNA hairpin located in the 5'UTR of the *cspA* mRNA. CspA interacts with a U-rich motif from the right arm of such hairpin. As a consequence, it melts the double-stranded structure avoiding RNase III cleavage and thus reducing CspA expression. The mechanism behind this auto-regulation depicted CspA as a putative antagonist of RNase III activity.

This Thesis shows how RNA chaperones, like *S. aureus* CspA, can specifically interact with RNA structures targeted by other RBPs and offers new ways of understanding CSP-mediated regulation. At the same time, it highlights the importance of intrinsic mRNA regulatory elements and proposes the interaction between them and RBPs as the key factor determining proper protein levels and, ultimately, allowing the correct development of organisms.

Sections of this Doctoral Thesis have been published in:
Nucleic Acids Research, 2018. DOI: 10.1093/nar/gkx1284
The regulon of the RNA chaperone CspA and its auto-regulation in *Staphylococcus aureus*. C. J. Caballero, P. Menendez-Gil, A. Catalan-Moreno, M. Vergara-Irigaray, B. García, V. Segura, N. Irurzun, M. Villanueva, I. Ruiz de los Mozos, C. Solano, I. Lasa and A. Toledo-Arana.

RESUMEN

RESUMEN

Las proteínas de unión a RNA (RBPs) son elementos esenciales para regular la expresión génica a nivel post-transcripcional. Los genomas codifican numerosas RBPs que incluyen diversos dominios proteicos. De forma general, estas pueden agruparse en proteínas ribosomales, ribonucleasas y chaperonas de RNA. Las *cold shock proteins* (CSPs) son un grupo de chaperonas de RNA, que poseen un dominio *cold shock* (CSD), y que se encuentran distribuidas en todos los seres vivos. A pesar de haber sido ampliamente estudiadas, aún se desconocen las dianas y la función biológica de la mayoría de ellas. El nombre de estas chaperonas se debe a que la expresión de la primera en ser identificada se inducía al bajar la temperatura de crecimiento. Sin embargo, el estrés por frío no actúa como activador en otros miembros de esta familia. Un solo genoma puede codificar varias CSPs que, al tener una elevada similitud de secuencia, ha llevado a plantear la existencia de una posible redundancia funcional entre ellas.

En esta Tesis hemos identificado y caracterizado el regulón, la especificidad, la función biológica y los mecanismos de regulación de la chaperona CspA utilizando *Staphylococcus aureus* como modelo. *S. aureus* es un patógeno de gran importancia clínica capaz de provocar una gran variedad de enfermedades y que, además, se ha convertido en un grave problema sanitario por la aparición de cepas multi-resistentes a los antibióticos. En el genoma de *S. aureus* existen tres variantes de CSPs con una identidad de secuencia superior al 70%. Para determinar si

existía una redundancia funcional entre las mismas, generamos un mutante en el gen *cspA*, que presentaba niveles reducidos de estafiloxantina (STX), el pigmento antioxidante que confiere el color característico a las colonias de *S. aureus*. Si los parálogos de CspA tuviesen la misma función, la producción de STX debería reestablecerse al expresar cualquiera de las tres variantes en dicha cepa. Sin embargo, los experimentos de complementación con un promotor heterólogo, revelaron que tan solo CspA restauraba la producción de STX. No obstante, a pesar de estar expresadas bajo el mismo promotor, observamos que los niveles de las proteínas CspB y CspC eran inferiores a los de CspA. Esto indicaba que las CSPs sufrían procesos de regulación post-transcripcional diferentes, lo que resultaba en concentraciones no comparables de las mismas. Con el fin de evitar esta limitación, modificamos genéticamente el mRNA de *cspA* para que codificase CspB o CspC. Estas mutaciones nos permitieron obtener una concentración similar para todas las CSPs. A pesar de ello, la producción de STX seguía siendo restaurada tan solo por CspA, lo que sugería una especialización funcional para cada una de las CSPs. Para determinar cuáles eran los motivos proteicos responsables de la especificidad, generamos construcciones quiméricas que intercambiaban regiones de CspB y CspC con CspA. De esta forma, descubrimos que las diferencias en los aminoácidos localizados en la mitad carboxi-terminal, y alejadas del dominio de unión a RNA, serían las responsables de la especificidad funcional de las CSPs.

Teniendo en cuenta que las CSPs de *S. aureus* resultaron no ser intercambiables entre sí, nos centramos en estudiar la función biológica específica de CspA. Para ello, combinando abordajes globales de proteómica cuantitativa e inmuno-precipitación de los complejos CspA-RNA, determinamos su regulón. Pudimos observar que CspA se comportaba como un regulador global en *S. aureus*, modulando la expresión de genes relacionados con el metabolismo de carbohidratos y ribonucleótidos, así como con la respuesta al estrés y virulencia. Como consecuencia, la deleción de *cspA* producía cambios fenotípicos evidentes como la desregulación en la formación del biofilm o la disminución de la resistencia a agentes oxidantes. Todo ello destacaba la importancia de CspA en la adaptación de *S. aureus* a condiciones de estrés.

El regulón mostraba que la unión de CspA a sus dianas podía afectar a la expresión proteica tanto positiva como negativamente. Generalmente, se asume que las CSPs son potenciadoras de la traducción al deshacer estructuras que dificultan el progreso de los ribosomas en los mRNAs. Por ello, la existencia de una regulación negativa fue una característica inesperada que nos llevó a estudiarla más en detalle. Descubrimos que la unión de CspA podía inhibir la expresión proteica mediante, al menos, dos mecanismos distintos. En el primer caso, CspA se unía a la 5'UTR del mRNA de *cspC,* disminuyendo la expresión de CspC. Esta unión no afectaba a los niveles del mRNA, por lo que cabía suponer que podía interferir con la traducción, probablemente dificultando el acceso del ribosoma. Los experimentos de traducción confirmaron que la sola

presencia de CspA era suficiente para inhibir la traducción de CspC *in vitro*. En el segundo caso, CspA reprimía su propia expresión al interferir con la acción de la endoribonucleasa III (RNase III). RNAse III potencia la traducción de CspA al procesar una horquilla de RNA localizada en la 5'UTR de su propio mRNA. CspA interaccionaba con dicha horquilla a través de un motivo rico en uridinas, provocando una desorganización de su estructura y evitando así el procesamiento por RNase III. Como consecuencia, la traducción de CspA se veía inhibida. Este mecanismo de autorregulación propone a CspA como un posible antagonista de la actividad de RNase III.

En esta Tesis, se pone de manifiesto que las chaperonas de RNA, como CspA, pueden interactuar de manera específica con estructuras de RNA, que a su vez pueden ser reconocidas por otras RBPs. Esto contribuye a un mejor entendimiento de la regulación mediada por este grupo de chaperonas de RNA. Además, se destaca la importancia de los elementos reguladores intrínsecos, presentes en cada mRNA, y se propone que la interacción de dichos elementos con distintas RBPs es un factor clave para la correcta expresión proteica que, en última instancia, permite el adecuado desarrollo de todos los seres vivos.

Secciones de esta Tesis Doctoral han sido publicadas en:

Nucleic Acids Research, 2018. DOI: 10.1093/nar/gkx1284
The regulon of the RNA chaperone CspA and its auto-regulation in *Staphylococcus aureus*
C. J. Caballero, P. Menendez-Gil, A. Catalan-Moreno, M. Vergara-Irigaray, B. García, V. Segura, N. Irurzun, M. Villanueva, I. Ruiz de los Mozos, C. Solano, I. Lasa and A. Toledo-Arana

INTRODUCTION

INTRODUCTION

Post-transcriptional regulation in bacteria

Control of gene expression is essential for bacteria as it economizes resources and adjusts the production of functional molecules to adapt to ever-changing environments. The central dogma of molecular biology states that a DNA molecule is transcribed into mRNA, which is then translated into protein. It was believed for many years that the levels of mRNA were a direct indicator of the amount of protein that would be produced (Crick, 1970). However, in the last decade different research groups have carried out investigations on regulation of gene expression that have come to question the role and importance of the mRNA within the genetic flow of information.

After transcription occurs, the amount of protein production is determined by different elements that modulate mRNA levels and ribosome accessibility. Most of these elements can positively or negatively modulate the expression of their mRNA targets. Such is the case of small RNAs (sRNAs), which are transcripts of about 50-300 nt long that act through base pairing with their mRNA targets (Waters and Storz, 2009; Storz *et al.*, 2011). In a similar manner, antisense RNAs (asRNA) base pair with perfect complementarity with mRNAs that are generated in the opposite

strand (Lasa *et al.*, 2012; Lasa and Villanueva, 2014). In addition, other RNA regulatory elements may be found within the mRNAs themselves. On the one hand, the 5' untranslated region (UTR) often contains regulatory structures like hairpins or riboswitches that can dictate the fate of gene expression (Kortmann and Narberhaus, 2012; Serganov and Nudler, 2013). On the other hand, the 3' UTRs that apart from containing the transcription terminator signal, also harbour regulatory elements capable of influencing gene expression by different mechanisms. Some of these include circularization of mRNA that blocks RBS, altering mRNA decay and/or binding proteins (Ruiz de Los Mozos *et al.*, 2013; López-Garrido *et al.*, 2014; Zhu *et al.*, 2015).

Finally, complex networks of RNA-protein interactions regulate gene expression at the post-transcriptional level. Eukaryotic cells encode hundreds of RNA binding proteins (RBPs) that play critical roles affecting RNA structures, interactions, biogenesis, modifications, turnover, localization, export and translation of mRNAs (Glisovic *et al.*, 2008). In bacteria, RBPs are numerous and contain diverse protein domains that recognize, bind, process or modify certain sequences and/or structures. As a consequence, bacterial RBPs could play diverse post-transcriptional roles, but elucidation of their precise molecular mechanisms require further investigations (Van Assche *et al.*, 2015). Regarding their functions, they can be classified into three generic groups: ribosomal proteins,

ribonucleases and RNA chaperones. In the following sections, the most relevant roles of these RBP groups are described.

Ribonucleases

Ribonucleases (RNases) are a group of enzymes that recognize and process RNA in function of their sequence and/or structure. Many RNases exist, each recognizing particular RNA patterns and displaying specific processing mechanisms.

RNase E is one of the most studied RNases, playing an essential role in Gram-negatives and processing the majority of its mRNAs (Ow and Kushner, 2002; Stead *et al.*, 2011; Hammarlöf *et al.*, 2015). It was first identified in *Escherichia coli* and found necessary for processing the 5S rRNA (Apirion and Lassar, 1978). The structure of RNase E involves two RNA binding domains that are important for target recognition, typically single stranded RNAs containing AU-rich regions instead of a particular nucleotide sequence (McDowall *et al.*, 1994). Despite being an endonuclease, its catalytic domain shows predilection for RNAs lacking the triphosphate cap at the 5'-end (Mackie, 1998). One of the major aspects of RNase E is that it is considered the scaffolding protein of the degradosome in Gram-negative bacteria, a protein complex that mediates bacterial RNA degradation. In addition to RNase E, it includes the

exonuclease PNPase, RNA helicase RhlB and a glycolytic enzyme known as enolase (Vanzo *et al.*, 1998; Carpousis, 2007). In Gram-positives, there are no RNase E homologs. However, they count with other RNases that somehow make up for it. Such is the case of RNase Y, an essential ribonuclease, which also targets single stranded AU-rich RNA regions. It has a preference for monophosphate 5'-ends and that affects a large amount of the mRNAs transcribed in *Bacillus subtilis* (Shahbabian *et al.*, 2009; Lehnik-Habrink *et al.*, 2011).

PNPase is a dual enzyme that combines the processing nature of an RNase with the synthetic capacity of an RNA polymerase. On the one hand, it degrades single stranded 3'-ends of at least 10-12 nt until it meets a stable secondary structure (Xu and Cohen, 1995; Coburn and Mackie, 1998). On the other hand, it generates A-tails in the same extremity, if the concentration of available NDPs is sufficient, which will ultimately affect the susceptibility of the entire RNA molecule to single stranded processing RNases like itself (Xu and Cohen, 1995; Mohanty and Kushner, 2000; Mohanty and Kushner, 2011). The importance of PNPase in bacterial systems resides in its central role in the degradosome. This is true for both Gram-negatives and Gram-positives. In the former, it associates itself with RNase E, helicase RhlB and enolase, while in the latter RNase Y, helicase CshA and a phosphofructokinase are its partners (Carpousis, 2007; Cho, 2017).

The 3'-exorribonuclease activity of PNPase is matched by that of RNase R and RNase II, two hydrolytic exonucleases that are also specialized in degrading single stranded 3'-ends (Nossal and Singer, 1968; Spickler and Mackie, 2000). RNase II initially seemed as functionally redundant with PNPase because, in addition to behaving as a 3'-exorribonuclease, it could not process double stranded structures of RNA and its deletion, alongside with that of PNPase, was unviable in *E. coli* (Donovan and Kushner, 1986; Spickler and Mackie, 2000). However, later studies reflected that both enzymes affected total RNA levels in the bacteria differently (Mohanty and Kushner, 2003). RNase R shows an additional role as it can manage through double stranded regions of RNA thanks to its unwinding capacity. Nevertheless, it needs the existence of a single stranded nucleotide stretch, like a poly-A tail, to recognize the substrate and initiate its activity (Cheng and Deutscher, 2005; Awano *et al.*, 2010).

The 5'-terminus of RNAs can also suffer processing from RNases. Homologs RNase J1 and RNase J2 were initially described as functional RNase E homologs in *B. subtilis* (Even *et al.*, 2005). Today, they are known for also degrading RNA molecules in a 5'-3' manner. RNase J1 seems to have a preference for 5'-monophosphorylated transcripts, which result from the activity of an RppH homologous in *B. subtilis.* (Mathy *et al.*, 2007; Richards *et al.*, 2011). RNase J1 and RNase J2, are encoded by *rnjA* and *rnjB* genes, respectively. Mutating both RNases to limit the expression of the former and delete the latter has a significant impact on

global gene expression. Although the number of transcripts whose levels are affecteds in RNase J1 and RNase J2 single mutants is very low, compared to a double mutant, RNAse J1 has a wider spectrum of targets than RNase J2. This is in agreement with the idea of RNase J1 having a more prominent exonuclease activity than RNase J2, in addition to being essential for cell viability (Mäder *et al.*, 2008; Mathy *et al.*, 2010).

RNase III plays an important role processing double stranded RNA configurations in bacteria. As showed by Lasa and colleagues, the impact of RNase III on the transcriptome of Gram-positive bacteria is vast (Lasa *et al.*, 2011). Up to three fourths of the annotated genome is affected by antisense transcription in *Staphylococcus aureus*, leading to the formation of overlapping transcripts. The repercussions of such double stranded RNA arrangements are the accumulation of short RNAs that result from the processing by this RNase (Lasa *et al.*, 2011). A later study, showed that a significant percentage of the annotated RNAs were immunoprecipitated alongside RNase III in *S. aureus*, confirming the broad influence of this ribonuclease (Lioliou *et al.*, 2012). Likewise, it was proposed that *E. coli* produces many antisense transcripts that generate double stranded RNA structures upon hybridizing with RNAs from the opposite strand. These double stranded RNA formations were isolated and found enriched in RNase III deficient mutants, indicating a processing by such RNase in a wild type strain (Lybecker *et al.*, 2014). These findings suggest that RNase III acts as a global regulator of antisense transcription,

processing overlapping coding sequences (CDSs), operons and UTRs (Lasa *et al.*, 2012). It also acts on hairpins that are formed as a result of internal base pairing in different RNA species. This can affect the translation of the resulting processed mRNA, as observed for the major cold shock protein A (CspA), the stability of the sRNA *rsaA* or the maturation of ribosomal RNA 16S in *S. aureus* (Lioliou *et al.*, 2012). Typically, these hairpins are 22 bp long, however, *in vitro* studies support that a minimum of 11 bp of double stranded RNA is required for efficient cleavage (Lamontagne and Elela, 2004; Gan *et al.*, 2006; Pertzev, 2006; Court *et al.*, 2013). Although no consensus sequence has been described for RNase III targets, certain nucleotides in the cleavage region are known to favour such processing more than others (Pertzev, 2006).

Ribosomal proteins

Ribosomal proteins are structural components of the ribosomes capable of binding RNA as well as interacting with each other. Since they are part of the ribonucleoprotein complex, they participate in translation of mRNAs.

The 30S or small subunit of the ribosome, which is responsible for recognizing and binding the mRNA throughout the translation process, contains 20 different proteins in addition to the 16S rRNA. The 50S or large ribosome subunit, is in charge of enabling the polymerization of the

new peptide during translation and contains around 30 proteins alongside the 23S and 5S rRNAs. In this case, its proteins do not interact with the translated mRNA. However, many of them are considered RNA binding proteins since they often bind rRNAs and/or act as feed-back negative regulators of their translation by interacting with their own mRNA leaders (Nashimoto and Nomura, 1970; Nomura, 1970).

Within the 30S, S1 protein is essential in a wide variety of bacteria with the exception of low GC content Gram-positives. It is composed of six S1 motifs (stretches of 72-75 amino acids), which are also present in many other RBPs and their role is to bind RNA (Bycroft *et al.*, 1997; Sorensen *et al.*, 1998; Salah *et al.*, 2009). S1 facilitates mRNA binding of the 30S about 11 nucleotides from the RBS (Sengupta *et al.*, 2001). For this reason, S1 plays a major role during the early stages of translation (Sorensen *et al.*, 1998; Delvillani *et al.*, 2011). Other than enabling ribosome binding to the mRNA, it holds an unwinding capacity, working as a chaperone that allows the ribosome to access the RBS of structured 5'UTRs of mRNAs. S1 would first bind to single stranded nucleotide stretches in close proximity to secondary structures and then melt them through a multistep process (Qu *et al.*, 2012; Mélodie Duval *et al.*, 2013).

The rest of the proteins are highly conserved among different bacterial species. Such is the case of S2, an essential protein that interacts with the mRNA during the formation of the Shine-Dalgarno (SD)-anti-SD duplex in early translation and recruits S1 into the ribosome by direct protein-protein

interaction (Yusupova *et al.*, 2006; Byrgazov *et al.*, 2012). S3 and S4 play a role in unwinding secondary structures of the mRNA substrate and promote ribosome progression, which in the case of S4 is linked to its anti-terminator properties (Torres *et al.*, 2001; Takyar *et al.*, 2005). S4 is also crucial during the initial steps of the 30S subunit assembly since it constitutes the first ribosomal protein that binds the 5' region of the 16S rRNA and thus forms a complex into which the remaining proteins are subsequently integrated (Kim *et al.*, 2014; Abeysirigunawardena *et al.*, 2017). S7 and S8 constitute two additional ribosomal proteins with RNA binding properties that interact with rRNA 16S (Saito and Nomura, 1994; Robert and Brakier-Gingras, 2001). S15 binds the 16S rRNA and serves as a bridge between the small and large ribosomal subunits (Culver *et al.*, 1999).

The 50S or large ribosomal subunit is a key player in the generation of the new polypeptide by accommodating amino acid loaded transport RNAs (tRNAs) within its structure. However, unlike the 30S, it does not bind the mRNA and, since it is a bigger complex, its composition is of about 30 proteins and two ribosomal RNAs, 23S and 5S.

L1, which is known for binding the 23S rRNA and forming a structure that is often referred as the "stalk" of the 50S, is involved in the translocation of tRNAs from the P (peptidyl) to the E (exit) site within the ribosome (Mikel Valle *et al.*, 2003). L4 also binds the 23S rRNA, contributing to its proper folding during ribosome assembly. Two additional elements of the 50S

subunit, L10 and L7/L12, form a complex of a pair of L7/L12 dimers that bind the C-terminal domain of L10 (Griaznova and Traut, 2000). Such complex recognizes the 23S rRNA, through L10, and recruits important translation factors (Wieden *et al.*, 2001; Iben and Draper, 2008). Lastly, L20 also binds the 23S rRNA thanks to its C-terminal domain (Raibaud *et al.*, 2003).

Most of the ribosomal proteins with RNA binding properties are known for their ability to down-regulate their own translation upon binding their mRNA. In many cases, they behave like RNA chaperones by blocking the RBS or the start codon but they can also act through other mechanisms. However, the mode of action for some of them still remains unknown (Sorensen *et al.*, 1998; Delvillani *et al.*, 2011).

RNA chaperones

RNA chaperones are proteins that bind RNA and have the ability to rearrange or fix their structure, affecting their stability, RNA accessibility and translation. RBPs that facilitate the base-pairing between sRNAs and their mRNA targets are also included in this group. In general, Hfq and CsrA are the most known RNA chaperones and, therefore, offer great examples of how this type of RBPs regulate gene expression of their targets.

Host factor Q, Hfq

Host factor Q, Hfq, was first described as a protein required for replication of phage QB in *E. coli* (Franze de Fernandez *et al.*, 1968). Hfq is widely known for acting as an RNA chaperone that facilitates and stabilizes interactions between sRNAs and mRNAs that do not necessarily present a complete base pairing complementation. The RNA chaperone activity is carried out by a ring-shaped complex from the union of six identical Hfq monomers, each comprising five antiparallel beta sheets and an alpha helix at the N-terminal. The first three beta sheets conform the Sm1 motif, while sheets 4 and 5 correspond to the Sm2 motif. Sm motifs are present in proteins of the Sm and LSm families, which are known for binding RNA (Schumacher *et al.*, 2002; Sauter *et al.*, 2003). The doughnut-shaped structure of Hfq can be divided in different RNA binding faces, each with a different preference for specific nucleotide sequences: the proximal face, the distal face and the rim (Updegrove *et al.*, 2016). The proximal face specializes in binding U-rich RNA motifs, like Rho-independent terminators. Removing the U-stretch from terminators of sRNAs may result in a deficient binding to Hfq. This was the case for SgrS sRNA that lost its ability to interact with Hfq and thus repress the expression of target mRNA *ptsG*, which encodes the major glucose transporter in *E. coli* (Otaka *et al.*, 2011; Ishikawa *et al.*, 2012). From a structural point of view, such U-rich

motifs are kept into pockets placed between contiguous monomers of the protein (Schumacher *et al.*, 2002; Sauer and Weichenrieder, 2011). Unlike the proximal face, the distal face recognizes sequences with high abundance of adenines (Link *et al.*, 2009; C Lorenz *et al.*, 2010). The position of the A-rich motif in the sRNAs and mRNAs seems of importance for Hfq activity. A good example to illustrate this is the mRNA that encodes the alternative sigma factor, *rpoS*, which in *E. coli* is targeted and positively regulated by sRNAs DsrA, RprA and ArcZ. These sRNAs interact with a hairpin present in the mRNA leader of *rpoS* that sequesters the RBS, releasing it and thus increasing translation. However, an A-rich motif located immediately upstream of the sRNAs binding site is required in the mRNA for Hfq to efficiently participate in the annealing and subsequent regulation of *rpoS* expression *in vivo* (Panja and Woodson, 2012; Peng, Soper, *et al.*, 2014).

The rim of Hfq contains a patch or arginines that is important for ensuring an optimal binding of Hfq to RNA targets. Removing these amino acids from the protein causes a significant decrease in Hfq binding efficiency. It is suggested that the rim is in some cases required as an additional binding site for the interaction with RNA targets. This is true for the *rpoS* mRNA, which binds the distal face and the rim of Hfq (Sauer *et al.*, 2012; Panja *et al.*, 2013; Peng, Curtis, *et al.*, 2014; Schu *et al.*, 2015).

In summary, the principal role of Hfq is to enhance RNA interactions. It acts as a platform where two different RNA molecules binding to different

faces of the protein meet. The interaction of two different RNAs with Hfq may involve structural rearrangements in which their complementary sequences are exposed and fall in close proximity to one another. As a consequence, gene expression and/or RNA stability are altered (Soper and Woodson, 2008; Beisel *et al.*, 2012; Panja and Woodson, 2012; Wagner, 2013).

Carbon storage regulator A, CsrA

Carbon storage regulator A (CsrA) is known as the regulator of the central carbon storage metabolism for its role in enhancing glycolysis and suppressing gluconeogenesis (Sabnis *et al.*, 1995). However, more recent studies highlight the global regulatory nature of this protein (Romeo *et al.*, 2013; Holmqvist *et al.*, 2016; Potts *et al.*, 2017). CsrA functions as a dimer, the interaction of two monomers leads to the formation of a β-barrel of 10 antiparallel β-sheets, which contain an RNA binding motif in each extreme, and two α-helixes hanging from each side of the protein (Gutiérrez *et al.*, 2005; Mercante *et al.*, 2006). It targets GGA motifs, typically found in the single stranded loops of short hairpins. These motifs are often found at the RBS (Dubey *et al.*, 2005; Holmqvist *et al.*, 2016). CsrA works as a post-transcriptional regulator of mRNAs at the level of translation. Most of the times it behaves as a repressor, occluding the SD

sequence and leading to early degradation of targeted mRNAs. Good examples of this are the *glgC* and *cstA* genes, responsible for glycogen synthesis and transport under carbon starvation in *E. coli*, respectively (Baker *et al.*, 2002; Dubey *et al.*, 2003). However, negative regulation does not always occur through the same mechanism. In the case of *sdiA* mRNA, the *N*-acylhomoserine-L-lactone receptor of *E. coli*, CsrA binds to the AUG codon preventing translation to initiate (Helen Yakhnin, Baker, *et al.*, 2011). In a more recent study, CsrA was found to repress the translation of IraD*,* a protein that prevents RpoS proteolysis. This was achieved by CsrA binding to the SD of the upstream ORF, located in the same operon (Park *et al.*, 2017).

In some cases, CsrA can enhance translation by stabilizing the mRNA or making it less susceptible to RNase processing. The mRNA of *flhDC*, the master regulator of flagellum synthesis, is protected from RNase E thanks to CsrA binding (Alexander V Yakhnin *et al.*, 2013).

The activity of CsrA is controlled by the interaction with CsrB and CsrC sRNAs. These sRNAs contain several CsrA-recognition motifs in their sequences. As a result, CsrB and CsrC act as sequestering factors, restraining the protein and preventing its action (Mu Ya Liu *et al.*, 1997; Weilbacher *et al.*, 2003; Babitzke and Romeo, 2007).

Cold shock proteins

Cold shock proteins (CSPs) are a group of RNA chaperones belonging to the cold shock domain (CSD) protein family which, as reflected in SMART database, is present in every kingdom of life (http://smart.embl.de/) (Letunic *et al.*, 2015). Besides being very widespread, a variable number of CSPs can be found within a single bacterial genome depending on the species (Graumann and Marahiel, 1998). For example, the Gram-negative *E. coli* and the Gram-positive *B. subtilis* that have been widely used as bacterial models to study CSPs, contain nine and three Csp-paralogs, respectively. The fact that CSPs present a high identity between them (above than 45%), that four of *E. coli* CSPs (CspA, CspB, CspG and CspI) are cold inducible and suggests a possible overlapping role (Nakashima *et al.*, 1996; Wang *et al.*, 1999; Xia *et al.*, 2001). A recent study in *Salmonella* has also shown that CspC and CspE have functional redundancy, supporting this notion (Michaux *et al.*, 2017). However, although certain CSPs may make up for the absence of others, complementation in some cases is non-existent, indicating specific roles for some CSPs (Graumann *et al.*, 1997; Xia *et al.*, 2001).

In conflict with their given name, several members of the CSP family are non-cold induced and their expression is either constitutive or activated upon different stresses other than cold. For this reason, it is thought that

CSPs might be required for bacterial adaptation to different environmental stresses. Thus, mutation of a specific CSP might prevent bacteria from adapting to cold, oxidative and osmotic stresses, as well as affect the intracellular lifestyle or stationary-growing phase (Willimsky *et al.*, 1992; Yamanaka *et al.*, 1998; Graumann and Marahiel, 1999; Schmid *et al.*, 2009; Brea D Duval *et al.*, 2010; Loepfe *et al.*, 2010; Czapski and Trun, 2014; Derman *et al.*, 2015). In *B. subtilis*, the presence of at least one *csp* gene is essential for viability (Graumann *et al.*, 1997). However, in other bacteria such as *Listeria monocytogenes,* none of its three CSPs are required for survival in non-stress growth conditions (Schmid *et al.*, 2009).

Throughout decades of research, different activities have been attributed to CSPs. Presently, it is believed that most of them act by melting RNA secondary structures in a low binding affinity manner, allowing ribosome progression and thus improving translation (Phadtare, 2004; Phadtare and Severinov, 2005; Phadtare and Severinov, 2010). Such unwinding capacity has also been proven to be a mechanism for transcription anti-termination of certain genes. This occurs in the *metY-rpsO* operon of *E. coli,* from which larger transcripts are generated upon CspE and CspC overexpression (Bae *et al.*, 2000; Phadtare, Tyagi, *et al.*, 2002; Phadtare *et al.*, 2003).

CSPs interact with RNA targets through hydrophobic interactions thanks to their CSD, which is about 7.4 kDa and consists of five anti-parallel β-sheets that form a β-barrel (Newkirk *et al.*, 1994; Schindelin *et al.*, 1994).

Two RNA binding motifs, RNAP1 and RNAP2, are located in β-sheets 2 and 3 respectively and are important for nucleic acid binding (Newkirk *et al.*, 1994), (Lee *et al.*, 2013). In *E. coli* and *B. subtilis*, mutations of residues from the aromatic cluster of CspA and CspB respectively, impaired nucleic acid binding (Schröder *et al.*, 1995; Hillier *et al.*, 1998). Several works on the structural features of this protein domain showed the key amino acids involved in the protein-nucleic acid interaction. The RNA binding and RNA melting activities are carried out by different amino acids. Thus, protein mutants in Phe^{17}, Phe^{30}, and His^{32} amino acids are unable to unwind RNA structures *in vitro* but retain the RNA binding capacity (Schröder *et al.*, 1995; Phadtare, Tyagi, *et al.*, 2002; Zeeb *et al.*, 2006; Sachs *et al.*, 2012). It is surprising that such a small protein domain carries out these functions in an ATP-independent manner, unlike the DEAD box helicases, which require ATP to rearrange RNA secondary structures (Redder *et al.*, 2015).

Auto-regulatory mechanisms of RNA-binding proteins.

Several RBPs have the ability to regulate their own expression. This suggests a generalized regulatory mechanism that makes it consistent to examine if the same kind of auto-regulation is true for the rest of the RBPs. Some examples of such mechanism are described below.

RNase III, can regulate its own expression through a negative feedback loop. Its mRNA secondary structure contains three double stranded loops that are located in the UTRs. The second hairpin, found in the 5'UTR, can be targeted and processed by RNase III itself and lead to the destabilization of the mRNA, ultimately affecting its translation (Matsunaga *et al.*, 1996).

Ribosomal protein S1 controls its own expression by binding the leader of its mRNA (Skouv *et al.*, 1990; Boni *et al.*, 2001). In addition, S1 may cooperate with S2 in down-regulating the translation of the latter and that of the elongation factor Ts. This is accomplished by interaction of S2 with the 5'UTR of its operon, *rspB-tsf*, in which genes *rspB* and *tsf* are translated into S2 and Ts proteins, respectively (Aseev *et al.*, 2008). S15, is able to bind its own mRNA and block its translation by stabilizing a pseudoknot secondary structure in the 5'UTR that, despite allowing 30S binding, avoids formation of an active ternary complex (Philippe *et al.*, 1993). Strikingly, such mechanism does not seem conserved, as indicated

in later observations in *Thermus thermophilus*. In this bacterium, S15 facilitates the formation of a secondary structure in its mRNA that masks the RBS, thus preventing ribosome recognition and translation (Serganov *et al.*, 2003).

Ribosomal proteins S4, S7, S8, L1 and L4 have a similar way of exercising their auto-regulation capacity. They bind either the mRNA leader or the start codon of the first gene of their own operon and repress, not only their own translation, but also that of proteins encoded in the same polycistronic transcript. This usually includes polymerase subunits, elongation factors and other ribosomal proteins (Jinks-Robertson and Nomura, 1982; Thomas *et al.*, 1987; Saito and Nomura, 1994; Robert and Brakier-Gingras, 2001; Mattheakis and Nomura, 1988; Stelzl *et al.*, 2003; Zengel *et al.*, 2003; Nevskaya *et al.*, 2005).

Hfq, recognizes two sequences present in its 5'UTR, one of them in the RBS. Binding both sites is necessary to prevent the initiation of the translational complex and inhibit its own expression (Većerek *et al.*, 2005).

The 5'UTR of *csrA* mRNA is structured in the form of four different hairpins that contain binding sites for CsrA in their loops. One of these target regions overlaps with the RBS, causing the chaperone to compete with the ribosome. When in sufficient amount, CsrA can prevent the ribosome from binding its mRNA and thus disrupt translation (Helen Yakhnin, Yakhnin, *et al.*, 2011).

In the case of CspA from *E. coli*, some evidences point out that it may be able to negatively regulate its protein expression by affecting mRNA stability. This concept emerged from the comparison of the expression and stability of the 5'UTR of CspA in a WT strain to those in Δ*cspA* (Bae *et al.*, 1997). In addition, a posterior study suggests that the existence of intrinsic regulatory elements in the 5'UTR could hinder translation of CspA {Yamanaka:1999vd}. Nonetheless, a more recent work pointed out that the whole *cspA* mRNA molecule secondary structure may behave as a regulator, transitioning between different forms in function of the temperature. The overall secondary structure at 4ºC would be less susceptible to RNases and more easily translated (Giuliodori *et al.*, 2010).

The RNA chaperone CspA in *S. aureus* as a working model

In this Thesis, we have focused on the identification and characterization of the biological function of the CSPs from *S. aureus*, a Gram-positive, ubiquitous, facultative anaerobic bacterium of major clinical importance. It can cause different kinds of diseases, from mild skin abscesses or impetigo to more serious endocarditis, pneumonia or bone infections (Tong *et al.*, 2015). Moreover, interest in studying this pathogen has increased with the emergence of strains capable of overcoming antibiotic treatment, also known as methicillin resistant *S. aureus* (MRSA) (Knox *et*

al., 2015; Tang *et al.*, 2017). For this reason, and the need to develop new antimicrobial strategies, it has been widely used as a model organism in molecular genetics. Research on the post-transcriptional regulation of gene expression in this bacterium has made noteworthy advancements in the past years. *S. aureus* produces the first described sRNA to have an effect on virulence, RNAIII (Novick *et al.*, 1993). Since then, many studies have been conducted, demonstrating the post-transcriptional implications of this sRNA in regulating expression of proteins that favour bacterial adhesion, cell wall remodelling, biofilm, capsule formation or genes related to toxin production (Bronesky *et al.*, 2016). *S. aureus* is also one of the first model organisms where pervasive antisense transcription was found, with RNase III playing a major role in processing the double-stranded structures that resulted from such overlapping transcription. The majority of the transcriptome appeared affected by this phenomena, revealing the presence of a genome-wide post-transcriptional regulatory mechanism (Lasa *et al.*, 2011; Lasa *et al.*, 2012). Likewise, the importance of 3'UTRs in bacterial gene regulation has been highlighted in *S. aureus*. This untranslated region has remained disregarded in bacteria, from a regulatory point of view, until recent times. However, it is widely accepted that 3'UTRs are essential elements controlling expression in eukaryotic cells by contributing to circularization of mRNAs. *S. aureus* holds the first described interaction between the 5' and 3'-UTRs of an mRNA in prokaryotes, which results in repression of its translation. It also serves as

one of the few examples in which the 3'UTR behaves as regulatory element by itself in bacteria (Ruiz de Los Mozos *et al.*, 2013). Although new examples of important regulatory RNAs are constantly being discovered in *S. aureus,* many remain to be characterized (Fechter *et al.*, 2014; Tomasini *et al.*, 2014; Bronsard *et al.*, 2017).

RNA-mediated regulatory mechanisms often require the participation of RBPs. However, this type of post-transcriptional regulators, such as RNA chaperones and their biological implications, remain uncharacterized in *S. aureus*. Here, we have concentrated our efforts in understanding the importance and mode of action of CspA, one of the three *S. aureus* CSPs, in regulating gene expression. CspA is one of the most abundant proteins present in the bacterial cytoplasm (Cordwell *et al.*, 2002), making it a common immunodominant antigen that appears during *S. aureus* human sepsis (Lorenz *et al.*, 2000). There is not much information available about this RBP but it has been suggested that the *cspA* gene could be slightly induced by cold stress (Katzif *et al.*, 2003). However, more recent results contradicted this possibility and highlighted *cspB* as the cold-shock inducible gene in *S. aureus* (Anderson *et al.*, 2006; Brea D Duval *et al.*, 2010; Lioliou *et al.*, 2012). A global protein interaction network study performed in this bacterium revealed that CspA would interact with at least 65 other proteins (Cherkasov *et al.*, 2011). Interestingly, a large portion of these interactions involved elements of the translation machinery, such as proteins from both ribosomal subunits and translation elongation factors.

This suggests a possible involvement of CspA in regulating translation (Cherkasov *et al.*, 2011). Up to date, the only known phenotype related to CspA is the production of staphyloxanthin (STX), the primary carotenoid pigment responsible for the characteristic golden yellow colour of *S. aureus* colonies (Katzif *et al.*, 2005).

The complex chromosomic organization surrounding *cspA* suggests that its expression is tightly regulated. It has been shown that its mRNA is synthetized from at least two different transcription start sites. The most active promoter is located in the intergenic region between *cspA* and the upstream gene SAOUHSC_01405 (also known as *msaA*), while the other is found immediately upstream of *msaA*. Consequently, *cspA* CDS could be translated from either a monocistronic or bicistronic transcript. An additional element in the genomic architecture surrounding *cspA* is an antisense transcript (*as-cspA*), complementary to the 5'UTR of the *cspA* monocistronic mRNA (Lioliou *et al.*, 2012; Sahukhal and Elasri, 2014; Koch *et al.*, 2014). Although the antisense mRNA is significantly less transcribed, the base pairing between both transcripts generates a double stranded RNA that can be processed by RNaseIII and could contribute to modulating the expression of CspA *in vivo* (Lioliou *et al.*, 2012). RNase III can also process a hairpin that is formed at the 5' UTR of the monocistronic *cspA* mRNA and generate a shorter version of the same mRNA. While both mRNA lengths can be translated, stability and toe-print experiments indicate that the generation of the short *cspA* mRNA makes

the transcript more stable and the RBS more accessible to the ribosome, respectively. Therefore, such mRNA length would likely be translated at a higher rate (Lioliou *et al.*, 2012).

OBJECTIVES

OBJECTIVES

Despite cold-shock proteins (CSPs) being important for the correct functioning of bacterial systems, little is known about their biological role. Details relative to CSPs-mediated regulatory mechanisms as well as the identification of CSPs RNA targets would provide an accurate understanding of their cellular function. Consequently, this Thesis was aimed at the functional characterization of the RNA chaperone CspA as a modulator of gene expression, using *S. aureus* as a model. The general objective of this work was divided into the following milestones:

1. Determine if the CSP paralogs have the same biological function in *Staphylococcus aureus*.
2. Identify the *in vivo* targetome map and the regulon of the RNA chaperone CspA.
3. Characterize the molecular mechanisms of the CspA-mediated regulatory pathways controlling gene expression.

MATERIAL AND METHODS

MATERIALS AND METHODS

Strains, plasmids, oligonucleotides and growth conditions

The bacterial strains, plasmids and oligonucleotides used in this study are listed in Tables 1, 2 and 3, respectively (these tables can be found at the end of this section). *S. aureus* strains were grown in Trypticase Soy Broth (Pronadisa) supplemented with 0.25% glucose (TSBg) or Mueller Hinton Broth (MH) (Sigma-Aldrich). *E. coli* was grown in LB broth (Pronadisa). B2 and SuperBroth media were used to prepare *S. aureus* and *E. coli* competent cells, respectively. For selective growth, media were supplemented with the appropriate antibiotics at the following concentrations: Erythromycin (Erm), 1.5 μg ml^{-1} or 10 μg ml^{-1}; Ampicillin (Amp), 100 μg ml^{-1}. In addition, 1.5 µM of cadmium was added when activation of the Pcad promoter was required.

Generation of mutant strains by homologous recombination

All mutant strains that were generated for the purpose of this study were obtained by marker-less homologous recombination, using the pMAD plasmid system (Arnaud *et al.*, 2004). Briefly, the CspA^{3xFLAG}, GdpP^{3xFLAG}

and Δ*cspA* strains were generated by a two-step procedure that replaces a chromosomic region by the corresponding mutant allele, contained in the pMAD plasmid (Jaione Valle *et al.*, 2003). The resulting modified strains were verified by PCR using the corresponding oligonucleotides, E and F (Table 3), and Sanger sequencing.

Total protein extraction

Preinocula were grown in 5 ml TSBg at 37ºC and 200 rpm overnight (ON). Bacterial concentrations were estimated by measuring their optical density (OD_{600}). Normalized bacterial aliquots were diluted 1:50 in 250 ml Erlenmeyer flasks containing 50 ml of TSBg. Cultures were incubated at 37ºC and 200 rpm and samples extracted at the experiment given time points. Samples were centrifuged for 3 min at 4,400 g and 4ºC. The resulting pellets were frozen in liquid nitrogen and stored at -80ºC until required. Pellets were thawed, washed once with PBS 1X and resuspended in 1ml of buffer containing 7 M Urea, 2 M Thiourea, 2% CHAPS and 50 mM Dithiothreitol (DTT) or 1 ml of phosphate buffered saline 1X (PBS) for LC-MS-based proteomics or Western blot experiments, respectively. Resuspended bacteria were transferred to Fast Prep tubes containing acid-washed 100 µm glass beads (Sigma) and mechanically lysed in a FastPrep-24 instrument (MP Biomedicals) for 45 s

and speed 6 twice. Tubes were centrifuged for 10 min at 21,000 g and 4ºC. Supernatants, containing total protein extracts, were quantified using Bradford protein assay kit (Bio-Rad) and samples prepared at the desired concentration in Laemmli buffer.

LC-MS-based comparative proteomics

Total protein samples (10 µg) were loaded and 1D SDS-PAGE was performed. Gels were fixed (50% methanol / 10% acetic acid), stained with Coomassie (Simply Blue Safe Stain, Invitrogen) and washed to reveal a unique band. The band was distained twice with 100 µl of Acetonitrile (AcN) at 40ºC for 5 minutes. In-gel tryptic digestion was performed using a 1:20 protein to trypsin ratio (Sequencing grade modified Trypsin-Promega) in 50 mM ammonium bicarbonate at 37ºC for 16 h, after a denaturation step with 10 mM DTT at 40ºC for 30 min, and an alkylation step with 25 mM Iodoacetamide at room temperature (RT) for 30 min. The resulting peptides were extracted with 1% formic acid (FA), 50% AcN. Peptide desalting, concentration and purification were performed using Pierce C18 Spint Tips, according to the manufacturer's instructions, and evaporated to dryness in a SpeedVac. Peptides were resuspended in 2% AcN and 0.1% FA prior to LC-MSMS analysis.

Peptide mixtures (1 µg) were separated by reverse phase chromatography using an Eksigent nanoLC ultra 2D pump fitted with a 75 µm ID column (Eksigent 0.075 × 250). Samples were first loaded for desalting and concentrating into a 0.5 cm length 100 µm ID precolumn packed with the same chemistry as the separating column. Mobile phases were 100% water 0.1% FA (buffer A) and 100% AcN 0.1% FA (buffer B). The column gradient was developed in a 240 min two step gradient from 5% buffer B to 25% buffer B in 210 min and 25% buffer B to 40% buffer B in 30 min. The column was equilibrated with 95% buffer B for 9 min and 5% buffer B for 14 min. During the whole process, the precolumn was in line with the column and the flow maintained all along the gradient at 300 nl min^{-1}. Eluted peptides from the column were analysed using a Sciex 5600 Triple-TOF system. Data was acquired upon a survey scan performed with the mass range set at 350–1250 m/z in a scan time of 250 ms. The top 35 peaks were selected for fragmentation. Minimum accumulation time for MS/MS was set to 100 ms, giving a total cycle time of 3.8 s.

MS/MS data acquisition was performed using Analyst 1.5.2 (Sciex) and spectra files were processed through Protein Pilot™ Software (v 5.0 Sciex); using Paragon™ Algorithm (Shilov *et al.*, 2007) for database search, Progroup™ for data grouping, and searched against the *S. aureus* NCTC 8325 proteome obtained from PATRIC (https://www.patricbrc.org/) (Wattam *et al.*, 2017). False discovery rate was performed using a non-lineal fitting method and a “result.group” file was created reporting only

results with 1% Global false discovery rate or better. Peptide quantification was performed using Progenesis LC–MS software (Nonlinear Dynamics). With the accurate mass measurements from full survey scans in the TOF detector and the observed retention times, runs were automatically aligned to compensate for between-run variations, and the quality of these alignments was manually supervised. Peptide identifications were exported from Protein Pilot to Progenesis LC–MS, where they were matched to the respective features. For quantification, only unique peptides were included, and the total cumulative abundance was calculated by adding up the individual abundance of all peptides assigned to each protein. One-way ANOVA was used to calculate the P-value based on the transformed values.

Gene functional analysis

Gene classification was performed following the SEED online database (http://pseed.theseed.org/) (Overbeek *et al.*, 2005). GO term enrichment analysis was performed by PANTHER Overrepresentation Test (release 20170413) using default parameters, which hierarchically sorted the results based on GO Ontology database released on 2017-06-26 (http://pantherdb.org) (Mi *et al.*, 2017).

Plasmid constructions

In general, plasmids used in this study were engineered by subcloning PCR fragments that were amplified from chromosomic DNA with DreamTaq DNA polymerase or Phusion High-Fidelity DNA Polymerase (Thermo Scientific) and the corresponding oligonucleotides (Table 3). The resulting PCR products were purified from agarose gels using NucleoSpin® Gel and PCR Clean-up Macherey-Nagel kit, ligated into the pJET 1.2 vector (Thermo Scientific) and cloned in *E. coli* XL1-Blue (Stratagene). Plasmids were purified from ON cultures with the NucleoSpin® Plasmid Macherey-Nagel kit and DNA fragments excised with FastDigest restriction enzymes (Thermo Scientific). The resulting DNA fragments were purified from agarose gels as described above and ligated, using the Rapid Ligation Kit (Thermo Scientific), into pCN51, pCN40 (Charpentier *et al.*, 2004) or pMAD (Arnaud *et al.*, 2004) plasmids. After verification by Sanger sequencing, the final constructs were introduced into *S. aureus* RN4220 by electroporation (Lee, 1995). Afterwards, the plasmids were purified from RN4220 strains and introduced into 15981 wild type and derivative strains by electroporation.

Specifically, pMAD plasmids, required for chromosomal deletion of *cspA* gene and 3xFLAG tagging of *cspA* and *gdpP* (Table 2), were generated by amplifying 400-500 nt of the corresponding flanking regions with

oligonucleotides AB and CD, respectively (Table 3). The resulting PCR fragments were cloned as described above.

pCspA^{3xFLAG} (Table 2) was constructed by amplifying a PCR product with the corresponding +1 and term oligonucleotides (Table 3) using chromosomic DNA from the 15981 *cspA*^{3xFLAG} strain (Table 1) as a template. DNA fragments were cloned into pCN51 using BamHI and EcoRI as described above.

pCspB^{3xFLAG} (Table 2) was generated by overlapping PCRs using oligonucleotides CspB +1 BamHI, 3xFcspB_B, 3xFcspB_C and CspC ter EcoRI (Table 3). The amplified product was ligated into pCN51.

pCspC^{3xFLAG} (Table 2) was generated by overlapping PCRs using oligonucleotides CspC +1 BamHI, 3xFcspC_B, 3xFcspC_C and CspC ter KpnI (Table 3). The amplified product was ligated into pCN51.

pCspA^{3xFLAG}ΔUTRs, pCspB^{3xFLAG}ΔUTRs and pCspC^{3xFLAG}ΔUTRs (Table 2) were constructed by amplifying PCR products with the forward CspA_D5UTR_BamHI, CspB_D5UTR_BamHI and CspC_D5UTR_BamHI oligonucleotides, respectively. CspX_D3UTR_KpnI was used as the reverse oligonucleotide (Table 3). The amplified products were ligated into pCN51.

Plasmids pCspA_orfB^{3xFLAG}, pCspA_orfC^{3xFLAG}, pCspA_MLB, pCspA_MLC, pCspAB^{3xFLAG}, pCspAC^{3xFLAG} (Table 2) were generated by subcloning synthetic DNA fragments into pCN51 from plasmids pMA-T_CspA_orfB^{3xFLAG}, pMA-T_CspA_orfC^{3xFLAG}, pMA-T_MLB, pMA-T_MLC,

pMA-T_CspAB^{3xFLAG} and pMA-T_CspAC^{3xFLAG}, respectively (Table 2). The latter were previously generated by GeneArt (Invitrogen, ThermoFisher Scientific) using BamHI and EcoRI sites.

pSigB^{3xFLAG} (Table 2) was constructed by amplifying two PCR products: (I) from pCN51 with oligonucleotides SpeI 3xF TT pCN51 and NarI 3xF TT pCN51, (II) from the chromosome with oligonucleotides SpeI SigB and EcoRI SigB (Table 3), which were cloned simultaneously in pCN40.

pCspA, pCspB and pCspC were generated by PCR with the corresponding +1 and term oligonucleotides (Table 3) using chromosomic DNA from the 15981 strain.

To create the pCN51 plasmid expressing the Δ5'UTR mutant (Table 2), oligonucleotides BamHI D5UTR cspA and CspA term EcoRI were designed to amplify a *cspA* mRNA lacking the 5'UTR while preserving the RBS (Table 3). pCN51 plasmids harbouring the mutated hairpin loop from *cspA* 5'UTR were constructed by subcloning synthetic DNA fragments from plasmids pMA-T_CspA_M5U and pMA-T_CspA_M5UC, previously generated by GeneArt (Invitrogen, ThermoFisher Scientific) using BamHI and EcoRI sites. Note that any gene cloned into the multiple cloning site of the pCN51 plasmid would be expressed as a chimeric operon, in which the first gene is *cadC*, the transcriptional regulatory protein that binds to the P*cad* promoter (Endo and Silver, 1995; Charpentier *et al.*, 2004).

Plasmid pGEX-6P-2::*cspA,* expressing the CspA protein fused to GST, was constructed by amplifying a PCR fragment with primers CspA-

GST_Fv and CspA-GST_Rv (Table 3) and subcloned into the pGEX-6P-2 vector (GE Healthcare Life Sciences), using BamHI and SalI sites.

Staphyloxanthin extraction and quantification

Staphyloxanthin (STX) extraction was performed as previously described with slight modifications (Pelz *et al.*, 2005). Preinocula were grown in tubes containing 5 ml MH at 37ºC and 200 rpm for 11 h. Bacterial concentrations were estimated by measuring their OD_{600}. Normalized bacterial aliquots were diluted 1:250 in 250 ml Erlenmeyer flasks containing 50 ml of MH. Cultures were incubated at 37ºC and 200 rpm for 15 h and centrifuged for 10 min at 4,400 g in pre-weighted 50 ml-Falcons. Bacterial pellets were washed with 50 ml of PBS and centrifuged again. After discarding the supernatants, bacterial pellets were weighted and resuspended in a variable volume of ethanol 96% (Merck) proportional to their weight. Bacterial suspensions (700 µl) were incubated at 45ºC for 2 h and centrifuged for 10 min at 21,000 g. Finally, the concentration of STX pigment contained in the supernatants, was determined by measuring optical density at 460 nm. The statistical comparison of STX levels means from biological triplicates was performed by a two-tailed paired *t*-test with a confidence interval of 95% using Prism software package (GraphPad).

Hydrogen peroxidase susceptibility assay

Hydrogen peroxidase (H_2O_2) assay was performed as previously described with slight modifications (Chia-I Liu *et al.*, 2008). Briefly, preinocula were grown in tubes containing 5 ml MH at 37ºC and 200 rpm for 10 h. Normalized bacterial preinocula were diluted 1:250 into 250 ml Erlenmeyer flasks containing 50 ml of MH. Cultures were grown at 37ºC and 200 rpm for 15 h. Bacterial density was adjusted to approximately $5x10^7$ CFU ml^{-1} and incubated at 37ºC and 200 rpm for 1 h in the presence or absence of H_2O_2 (final concentration of 0.09%). Bacterial viability was addressed by plating serial dilutions on TSA plates and incubating them at 37ºC ON.

Western Blotting

Total protein samples were boiled at 100ºC for 5 min, loaded and run into 12% SDS-polyacrylamide gels. Resolved proteins were transferred to nitrocellulose membranes (Amersham Biosciences) or stained with Coomassie brilliant blue R250 (Sigma) for Western Blot and loading control respectively. Membranes were treated for at least 1 h at RT with a blocking solution (5% skimmed milk powder, 0.1% Tween 20 PBS). After

washing several times with 0.1% Tween 20 PBS, membranes were incubated for 1.5 h at RT with anti-FLAG antibodies (Sigma) diluted 1:1000 in the blocking solution. Several washing steps for a period of 45 min were performed and the flagged proteins were developed using the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific), following the manufacturer instructions.

RIP-chip analysis

RNA-Binding Protein Immunoprecipitation-Microarray (RIP-chip) was performed as previously described with some modifications (Jain *et al.*, 2010). Preinocula were grown in 5 ml TSBg at 37ºC and 200 rpm ON. Bacterial concentrations were estimated by measuring their OD_{600}. Subsequently, bacteria were diluted 1:100 into 500 ml Erlenmeyer flasks containing 250 ml of TSBg. Cultures were incubated at 37ºC and 200 rpm until mid-exponential phase was reached. Next, bacteria were transferred to 50 ml falcon tubes containing 1.4 ml of 37% formaldehyde solution (Sigma F8775). The tubes were incubated at RT for 15 min with occasional inversion for mixing. Sterile 2.5 M glycine (5 ml) was added to each tube to stop the crosslinking reaction and incubated at RT for 5 min. Then, falcon tubes were centrifuged at 3,500 g and 4ºC for 6 min. Pellets were washed with 50 ml cold Tris-buffered saline (150 mM NaCl, 10 mM

Tris-HCl, pH 7.5), centrifuged and frozen at -80ºC ON. Pellets were resuspended in 0.5 ml of lysis buffer 1 (10 mM Tris-HCl, pH 8.0, 20% sucrose, 50 mM NaCl, 10 mM EDTA) containing protease inhibitors (Roche), RNAsin 40 U ml^{-1} (Promega) and 10 µl of lysostaphin 10 mg ml^{-1} (Sigma) and incubated at 37ºC for 15 min. Cell lysis was completed by adding 1.5 ml of cold lysis buffer 2 (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS), containing protease inhibitors and RNAsin. Cell debris was removed from the samples by centrifugation at 21,000 g and 4ºC for 15 min. Supernatants (1.6 ml) were pre-incubated with 100 µl of protein G-Sepharose beads (Pierce) at 4ºC for 1 h. Prior to use, beads were washed 4 times with 250 µl of lysis buffer and stored in the same buffer, including protease inhibitors and RNAsin. To immunoprecipitate cross-linked RNA complexes, the supernatants were retrieved and incubated with G-Sepharose beads conjugated with anti-Flag antibody (Sigma) (180 µl of G-sepharose beads and 11 µl of anti-Flag antibody incubated ON) at 4ºC for 2 h on a rotating wheel. Complexes were pulled down using Spin-X columns (CORN 8160) and washed several times. The first wash was performed in 1 ml of cold lysis buffer supplemented with 500 mM NaCl. The second wash, with 1 ml of cold wash buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% Nonidet P-40, 0.5% sodium deoxycholate). Finally, they were washed with 1 ml Tris-EDTA pH 7.5 and RNAsin at 4ºC. The beads, contained in the last solution, were then

transferred to a 2 ml Eppendorf, centrifuged, and resuspended in 150 µl of elution buffer (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1% SDS) supplemented with RNAsin. For reverse cross-linking, samples were incubated at 65ºC for 1 hour and treated with acid phenol and chloroform. The RNAs contained in the aqueous phase were precipitated using 5 µl of glycogen Ambion (AM 9510), 25 µl 3M sodium acetate and 625 µl of ethanol at -80ºC ON. The pellets were retrieved by centrifugation at 21,000 g and 4ºC for 30 min. RNAs were resuspended in DNase buffer including RNAsin and treated with TURBO DNase (Ambion) at 37ºC for 30 min. After phenol-chloroform extractions RNAs were precipitated as indicated above and resuspended in 20 µl of DEPC H_2O. 1 µl of each RNA sample was loaded into Agilent RNA Nano LabChips (Agilent Technologies) to determine their concentration and quality. The rest of the RNA sample was used as a template for cDNA synthesis, following the recommendations of Affymetrix protocol and hybridized on Affymetrix custom *S. aureus* tiling microarrays, as previously described (Segura *et al.*, 2012). CspA-binding signals were normalized with Tiling Array Software (TAS) to generate signal intensity files that we then loaded into the Integrated Genome Browser (IGB) using the *S. aureus* NCTC 8325 genome as a reference (Iandolo *et al.*, 2002; Freese *et al.*, 2016).

CspA binding peak calling

The peak calling was performed by two complementary methods. On the one hand, peak signals above the background were extracted using IGB by applying a threshold value. Prior to establishing the threshold value, the distribution of intensity signals from the 363,127 probes contained in the tiling microarrays was calculated for each of the RIP-chip samples. The CspA^{3xFLAG}, GdpP^{3xFLAG}, and WT RIP-chip samples had background noise profiles with a mean signal intensity of 5.5 ± 0.7, 5.4 ± 0.3 and 5.0 ± 0.3 respectively. Since CspA^{3xFLAG} and GdpP^{3xFLAG} RIP-chips had similar signal noise, the background threshold was defined as 5.4 (the mean signal intensity of the 3xFLAG-negative control) plus four standard deviations. Considering this threshold value, a total of 18,222 probes were detected above the background signal in CspA^{3xFLAG} but 2,037 and 1,847 of them were also present in the negative control samples, WT and GdpP^{3xFLAG}, respectively. Most of the signals above this threshold in the WT and GdpP^{3xFLAG} RIP-chip samples (95.6% and 92.3%, respectively) corresponded to ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs). Therefore, these signals were considered non-specific and excluded from subsequent analyses. In addition, peaks corresponding to repeated regions, which could give false positives, were not considered either (Waldminghaus and Skarstad, 2010). On the other hand, CisGenome, an

integrated software for processing raw microarray that defines peaks thanks to computing false discovery rates (FDR), was used. Recommended normalization parameters for Affymetrix tiling microarrays were applied (Ji *et al.*, 2011). The peak calling results and normalized RIP-chip signals are fully available at our RNA map web browser (http://rnamaps.unavarra.es/).

RNA extraction and Northern blotting

Bacteria were grown as described for total protein extraction (see above). Cultures were centrifuged for 3 min at 4,400 g and 4ºC. Pellets were then frozen in liquid nitrogen and stored at -80ºC ON. RNA extractions were performed as previously described (Toledo-Arana *et al.*, 2009). An appropriate amount of RNAs were mixed with formaldehyde loading dye (Ambion), denatured for 5 min at 65ºC and run in 1.25 % agarose gels. Gels were submerged in an ethidium bromide solution and RNA integrity and loading were verified by exposure to UV light. RNAs were then transferred to 0.2 µm pore size Nitran N membranes (GE Healthcare Life Sciences) by capillarity, using NorthernMax Transfer Buffer (Ambion), for 1.5 h at RT. The transferred RNAs were crosslinked to the membranes by exposing them to UV light inside a UV Stratalinker 1800 (Stratagene). Membranes were then placed into hybridization tubes and prehybridized

with ULTRAHyb solution (Ambion) for at least 30 min at 40ºC (for oligonucleotide probe) inside a rotating oven. After the prehybridization step, the corresponding radioactively labelled oligonucleotide was added and incubated ON. Membranes were washed three times for 5 min by addition of preheated 2X SSC, 0.1% SDS at 40ºC followed by several washes with 0,2xSSC, 0,1% SDS at RT until background signal was eliminated. Membranes were developed by autoradiography for different time periods.

Purification of recombinant CspA protein

To produce recombinant CspA protein, *E. coli* BL21 strain carrying pGEX-6P-2::cspA plasmid was grown in LB medium supplemented with Amp 100 µg ml^{-1} at 37°C and 250 rpm until an OD_{600} of 0.5 was reached. Isopropyl-D-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.4 mM and cell growth resumed for 5 h under the same conditions. Cells were collected by centrifugation at 5,000 g for 30 min, resuspended in phosphate buffer supplemented with lysozyme (1 mg ml^{-1}; Sigma) and incubated for 30 min at RT. Next, bacteria were lysed by sonication and centrifuged for 20 min at 16,000 g and 4ºC. The soluble fraction was incubated on ice for 30 min in the presence of DNase and RNase (10 µg ml^{-1}; Thermo Scientific). The GST-CspA fusion protein was purified from

clarified lysates using a GSTrap FF 5-ml column and an AKTAprime plus chromatography system (GE Healthcare Life Sciences). For PreScission Protease on-column cleavage, 10 U of protease mg^{-1} of bound GST fusion protein was used. To achieve the highest purity, size exclusion chromatography was applied with a HiPrep 16/60 Sephacryl S-100 HR column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 7.4, containing 500 mM NaCl, and run with a flow rate of 0.5 ml min^{-1}. Fractions containing the protein of interest were pooled and buffer exchanged by dialysis against ice cold 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 50 mM potassium chloride, and 10% glycerol using a Slide-A-Lyzer Dialysis Cassette (ThermoFisher Scientific) with a molecular mass cut-off of 3.5 kDa. Protein purity was determined by SDS-PAGE analysis and samples were quantified using the Bio-Rad protein assay (Bio-Rad) and stored at -20°C.

In vitro transcription

PCR fragments containing the T7 promoter were amplified using the oligonucleotides listed in Table 3. Note that Fw primer carried the T7-RNA Polymerase binding site at the 5'-end. *In vitro* transcription was carried out using the MEGAscript T7 transcription kit (Ambion) following the manufacturer´s recommendations. Transcripts were run in a denaturing

6% polyacrylamide gel and visualized with UV-light after staining with Ethidium Bromide. Bands were cut and incubated with 500 µl of ammonium acetate pH 5.5 2M at 15ºC for 2 h. Next, 500 µl of acid phenol pH 4.5 were added to the samples before mixing them ON at 15ºC. After a centrifugation for 5 min at 21,000 g and 4ºC, the aqueous phase was transferred to a tube containing 500 µl of acid phenol. The aqueous phase was again retrieved by centrifugation and washed with 600 µl of chloroform. RNA from the aqueous phase was precipitated with 1/10 volume of 3M sodium acetate pH 4.5 and 3 volumes of 96% ethanol at -80ºC for at least 1 h. Samples were then centrifuged at 21,000 g for 30 min and 4ºC and the resulting pellets washed with 70% ethanol. Pellets were air-dried and resuspended in the desired DEPC H_2O (Ambion). RNA integrity was checked through agarose gel electrophoresis.

5'-end labelling of RNA and ssDNA oligonucleotides

Before labelling, RNA and ssDNA oligonucleotides were dephosphorylated with FastAP (Thermo Scientific) at 37ºC for 1h. The enzyme was inactivated at 75ºC for 5 min. Dephosphorylated nucleic acids were incubated with ^{32}P-γ-ATP (Perkin-Elmer) and PNK (Thermo Scientific) for 30 min at 37ºC. EDTA was added and the enzyme was inactivated for 15

min at 75ºC. Labelled nucleic acids were purified with Illustra MicroSpin G-50 columns (GE Healthcare) and stored at -20ºC until used.

Electrophoretic mobility shift assays, EMSAs

Labelled ssDNA were diluted at the appropriate concentration by adding DEPC H_2O and 5X renaturing buffer (100 mM K-HEPES pH7.5, 50 mM $MgCl_2$, 250 mM KCl). Samples were denatured at 90ºC for 1 min, chilled on ice 1 min and renatured at 37ºC for 15 min. Labelled samples were mixed with RiboLock RNase Inhibitor (Thermo Scientific), 2X reaction buffer (20 mM Tris-HCl pH 7.5, 60 mM KCl, 40 mM NH_4Cl, 3 mM DTT, 0,02 mg ml^{-1} BSA, 10 mM $MgCl_2$) and increasing concentrations of purified recombinant CspA protein. The mixture was incubated at 37ºC for 15 min. Loading dye (50 % glycerol and bromophenol blue) was added to all samples and these were loaded into a non-denaturing 6 or 10% polyacrylamide gel (pre-run for 10-20 min) in 0.5X TBE at 110V, 4ºC for 1-5h. The gel was then dried and developed by autoradiography for different time periods. Dissociation constants (Kd) were calculated from signal intensities of unbound ssDNA oligonucleotides.

In vitro translation assay

In vitro translation assay was performed using the PURExpress® *In Vitro* Protein Synthesis kit (New England Biolabs) following the manufacturer's recommendations. Briefly, 0.06 pmoles of RNA were mixed with 10 µl of solution A, 7.5 µl of solution B, 20 U of Ribolock (Thermo) and 2.5 nmoles of BSA or CspA accordingly. Samples were incubated for 4 h at 37 ºC and then loaded into 12% SDS-polyacrylamide gels for Western blotting as described above.

FAM-quencher assay

The ssDNA oligonucleotide labelled with the 6-FAM and the Black Hole Quencher (BHQ_1) molecules in the 5' and 3' extremity, respectively, were acquired from Integrated DNA Technologies. As a test control, 1 pmol of the oligos were mixed with 12.5 µl of 1X CspA storage buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 50 mM potassium chloride) and 2.5 µl of 10X reaction buffer (100 mM Tris-HCl pH 7.5, 300 mM KCl, 200 mM NH_4Cl, 15 mM DTT, 50 mM $MgCl_2$) in a final volume of 25 µl. FAM fluorescence was then measured at 5 different temperatures (25ºC, 37ºC, 45ºC, 55ºC and 65ºC) using the Aria Mx Real-Time PCR System (Agilent

Technologies). Experiments were performed by diluting 1 pmol of the oligos in 10X Reaction buffer and 4U of Ribolock (Thermo) to a final volume of 30 µl. Next, 7 nmol of CspA or BSA were added to the mix accordingly and incubated at 37ºC for 15 min. Subsequently, 10 µl of Proteinase K (20 mgl ml^{-1}) (Sigma) were added and the mix was incubated again for 30 min at 37ºC. The samples were then incubated for 5 min at 65ºC. FAM emission was registered throughout all the incubation steps.

Table 1. Strains used in this study

Strains	Relevant characteristic(s)	BGR ID[a]	Source or reference
Staphylococcus aureus			
15981	Wild type (WT) strain. MSSA clinical isolate; biofilm positive; PNAG-dependent biofilm matrix	8	(Jaione Valle *et al.*, 2003)
Δ*cspA*	15981 strain with deletion of the *cspA* gene	83	This study
15981 *cspA*^{3xFLAG}	15981 strain expressing the 3xFLAG-tagged CspA protein from the chromosome	239	This study
15981 *gdpP*^{3xFLAG}	15981 strain expressing the 3xFLAG-tagged GdpP protein from the chromosome	305	This study
WT pCN51	15981 strain carrying the pCN51 plasmid	105	This study
Δ*cspA* pCN51	15981 Δ*cspA* strain carrying the pCN51 plasmid	107	This study
Δ*cspA* pCspA	15981 Δ*cspA* strain carrying the pCspA plasmid	108	This study
Δ*cspA* pCspB	15981 Δ*cspA* strain carrying the pCspB plasmid	109	This study
Δ*cspA* pCspC	15981 Δ*cspA* strain carrying the pCspC plasmid.	110	This study
Δ*cspA* pCspA^{3xFLAG}	15981 Δ*cspA* strain carrying the pCspA^{3xFLAG} plasmid	268	This study
Δ*cspA* pCspB^{3xFLAG}	15981 Δ*cspA* strain carrying the pCspB^{3xFLAG} plasmid	351	This study
Δ*cspA* pCspC^{3xFLAG}	15981 Δ*cspA* strain carrying the pCspC ^{3xFLAG} plasmid.	652	This study
Δ*cspA* pCspA^{3xFLAG}ΔUTRs	15981 Δ*cspA* strain carrying the pCspA^{3xFLAG}ΔUTRs plasmid	436	This study
Δ*cspA* pCspB^{3xFLAG}ΔUTRs	15981 Δ*cspA* strain carrying the pCspB^{3xFLAG}ΔUTRs plasmid	437	This study
Δ*cspA* pCspC^{3xFLAG}ΔUTRs	15981 Δ*cspA* strain carrying the pCspC^{3xFLAG}ΔUTRs plasmid.	438	This study
Δ*cspA* pCspA_orfB^{3xFLAG}	15981 Δ*cspA* strain carrying the pCspA_orfB^{3xFLAG} plasmid	763	This study
Δ*cspA* pCspA_orfC^{3xFLAG}	15981 Δ*cspA* strain carrying the pCspA_orfC^{3xFLAG} plasmid	761	This study
Δ*cspA* pCspA_MLB	15981 Δ*cspA* strain carrying the pCspA_MLB plasmid	696	This study
Δ*cspA* pCspA_MLC	15981 Δ*cspA* strain carrying the pCspA_MLC plasmid	698	This study
Δ*cspA* pCspA_B^{3xFLAG}	15981 Δ*cspA* strain carrying the pCspA_B^{3xFLAG} plasmid	760	This study
Δ*cspA* pCspA_C^{3xFLAG}	15981 Δ*cspA* strain carrying the pCspA_C^{3xFLAG} plasmid	762	This study
WT pSigB$^{3x\ FLAG}$	15981 strain carrying the pSigB^{3xFLAG} plasmid	656	This study
WT pCspC^{3xFLAG}	15981 strain carrying the pCspC^{3xFLAG} plasmid.	651	This study
WT pCspA^{3xFLAG}	15981 strain carrying the pCspA^{FLAG} plasmid	505	This study
WT pCspA^{3xFLAG}Δ5'UTR	15981 strain carrying the pΔ5'CspA^{3xFLAG} plasmid	649	This study
WT pM5U^{3xFLAG}	15981 strain carrying the pM5U^{3xFLAG} plasmid	746	This study
WT pM5UC^{3xFLAG}	15981 strain carrying the pM5UC^{3xFLAG} plasmid	748	This study
Δ*cspA* pSigB$^{3x\ FLAG}$	15981 Δ*cspA* strain carrying the pSigB$^{3x\ FLAG}$ plasmid	657	This study
Δ*cspA* pCspA^{3xFLAG}Δ5'UTR	15981 Δ*cspA* strain carrying the pΔ5'CspA^{3xFLAG} plasmid	650	This study
Δ*cspA* pM5U^{3xFLAG}	15981 Δ*cspA* strain carrying the pM5U^{3xFLAG} plasmid	747	This study

Continued in the following page

Table 1. Continued

Δ*cspA* pM5UC^{3xFLAG}	15981 Δ*cspA* strain carrying the pM5UC^{3xFLAG} plasmid	749	This study
Δ*rnc*	15981 strain with deletion of the *rnc* gene	409	(Lasa *et al.*, 2011)
Δ*rnc* pCspA^{3xFLAG}	15981 Δ*rnc* strain carrying the pCspA^{3xFLAG} plasmid	1036	This study
Δ*rnc* Δ*cspA*	15981 strain with deletion of the *rnc* and *cspA* genes	989	This study
Δ*rnc* Δ*cspA* pCspA^{3xFLAG}	15981 Δ*cspA* Δ*rnc* strain carrying the pCspA^{FLAG} plasmid	968	This study
Escherichia coli			
XL1-Blue	Strain used for cloning experiments	1	Stratagene
XL1-Blue pCN40	XL1Blue strain carrying the pCN40 plasmid	18	(Charpentier *et al.*, 2004)
XL1-Blue pCN51	XL1Blue strain carrying the pCN51 plasmid	20	(Charpentier *et al.*, 2004)
BL21pGEX-6P-2::cspA	BL21 (DE3) expressing CspA with a GST tail from the pGEX-6P plasmid	631	This study

[a] Identification number of the strains stored at the Laboratory of Bacterial Gene Regulation.

Table 2. Plasmids used in this study

Plasmids	Relevant characteristic(s)	Source or reference
pMAD	*E. coli-S. aureus* shuttle vector with a thermosensitive origin of replication for Gram-positive bacteria. It contains the *bgaB* gene encoding a β-galactosidase under the control of a constitutive promoter as reporter of plasmid presence. Amp^R, Erm^R	(Arnaud *et al.*, 2004)
pCN40	*E. coli-S. aureus* shuttle vector to express genes under the control of the P*blaZ* promoter. Low copy number. Amp^R-Em^R	(Charpentier *et al.*, 2004)
pCN51	*E. coli-S. aureus* shuttle vector to express genes under the control of the P*cad* inducible promoter. Low copy number. Amp^R-Em^R	(Charpentier *et al.*, 2004)
pMAD *cspA*	pMAD plasmid containing the allele for deletion of the *cspA* coding region	This study
pCspA	pCN51 plasmid expressing the CspA protein from the entire *cspA* mRNA molecule	This study
pCspB	pCN51 plasmid expressing the CspB protein from the entire *cspA* mRNA molecule	This study
pCspC	pCN51 plasmid expressing the CspC protein from the entire *cspA* mRNA molecule	This study
$pCspA^{3xFLAG}$	pCN51 plasmid expressing the 3xFLAG-tagged CspA protein from the entire *cspA* mRNA molecule	This study
$pCspB^{3xFLAG}$	pCN51 plasmid expressing the 3xFLAG-tagged CspB protein from the entire *cspB* mRNA molecule	This study
$pCspC^{3xFLAG}$	pCN51 plasmid expressing the 3xFLAG-tagged CspC protein from the entire *cspC* mRNA molecule	This study
$pCspA^{3xFLAG}$ΔUTRs	pCN51 plasmid expressing the 3xFLAG-tagged CspA protein from a *cspA* mRNA molecule lacking the UTRs while preserving the RBS and transcriptional terminator	This study
$pCspB^{3xFLAG}$ΔUTRs	pCN51 plasmid expressing the 3xFLAG-tagged CspB protein from a *cspB* mRNA molecule lacking the UTRs and containing the RBS and transcriptional terminator from *cspA*	This study
$pCspC^{3xFLAG}$ΔUTRs	pCN51 plasmid expressing the 3xFLAG-tagged CspC protein from a *cspC* mRNA molecule lacking the UTRs and containing the RBS and transcriptional terminator from *cspA*	This study
pMA-T_CspA_orfB^{3xFLAG}	pMA-T plasmid carrying the synthetic CspA_orfB^{3xFLAG}	This study
pMA-T_CspA_orfC^{3xFLAG}	pMA-T plasmid carrying the synthetic CspA_orfC^{3xFLAG}	This study
pCspA_orfB^{3xFLAG}	pCN51 plasmid expressing a chimeric *cspA* mRNA with nucleotide substitutions to codify CspB	This study
pCspA_orfC^{3xFLAG}	pCN51 plasmid expressing a chimeric *cspA* mRNA with nucleotide substitutions to codify CspC	This study
pMA-T_MLB	pMA-T plasmid carrying the synthetic MLB gene	This study
pMA-T_MLC	pMA-T plasmid carrying the synthetic MLC gene	This study
pCspA_MLB	pCN51 plasmid expressing a chimeric *cspA* mRNA with CspB RNP1 and RNP2 connecting loop	This study
pCspA_MLC	pCN51 plasmid expressing a chimeric *cspA* mRNA with CspC RNP1 and RNP2 connecting loop	This study
pMA-T_CspAB^{3xFLAG}	pMA-T plasmid carrying the synthetic CspA_B^{3xFLAG} gene	This study
pMA-T_CspAC^{3xFLAG}	pMA-T plasmid carrying the synthetic CspA_B^{3xFLAG} gene	This study

Continued in the following page

Table 2. Continued

pCspAB^{3xFLAG}	pCN51 plasmid expressing a chimeric *cspA* mRNA with nucleotide substitutions to codify the carboxi-half of CspB	This study
pCspAC^{3xFLAG}	pCN51 plasmid expressing a chimeric *cspA* mRNA with nucleotide substitutions to codify the carboxi-half of CspC	This study
pMAD *cspA*^{3xFLAG}	pMAD plasmid containing the allele for insertion of the 3xFLAG at the C-terminus of the CspA protein	This study
pMAD *gdpP*^{3xFLAG}	pMAD plasmid containing the allele for insertion of the 3xFLAG at the C-terminus of the GdpP protein	This study
pSigB^{3xFLAG}	pCN40 plasmid expressing the 3xFLAG-tagged SigB protein	This study
pΔ5'CspA^{3xFLAG}	pCN51 plasmid expressing the 3xFLAG-tagged CspA protein from a *cspA* mRNA lacking the 5'UTR	This study
pMA-T_CspA_M5U	pMA-T plasmid carrying the synthetic *cspA_M5U* gene	This study
pMA-T_CspA_M5UC	pMA-T plasmid carrying the synthetic *cspA_M5UC* gene	This study
pCspA^{3xFLAG}-M5U	pCN51 plasmid expressing the 3xFLAG-tagged CspA protein from a *cspA* mRNA carrying a mutation that substituted the 5U motif within the 5'UTR	This study
pCspA^{3xFLAG}-M5UC	pCN51 plasmid expressing the 3xFLAG-tagged CspA protein from a *cspA* mRNA carrying a mutation that substituted the 5U motif and a compensatory one to restore the stem-loop structure within the 5'UTR	This study

Table 3. Oligonucleotides used in this study

Oligonucleotide name	Sequence [a]
Probe for Northern blots assays	
anti_3xFLAG_probe	TTTATCGTCGTCATCTTTGTAGTCGATATCATGATCTTTATAATCACCGTCATG GTCTTTGTAGTC
Construction of pMAD plasmid for deletion of *cspA* gene	
CspA_A_BamHI	*GGATCC*TCAAACTTGATTTTCAGAGG
CspA_B_SalI	*GTCGAC*AATCTGAAACCTCCAAGACT
CspA_C_SalI	*GTCGAC*ATTCTTAGATTTGAATCATTG
CspA_D_EcoRI	*GAATTC*GATTTTCTAATTCCATCATCTCTG
CspA_E	TGTCTCATTTTTACCACCTC
CspA_F	CCACACTTTAAACAAGCATC
Construction of pMAD plasmid for chromosomic 3xFLAG-labeling of the *cspA* gene	
CspA_3xF_A_EcoRI	*GAATTC*ACTACCAATAATTGATTAAATGAATATT
CspA_3xF_B_NcoI	*CCATGG*TTACTATTTATCGTCGTCATCTTTGTAGTCGATATCATGATCTTTATA ATCACCGTCATGGTCTTTGTAGTCTAGTTTAACAACGTTTGCAGC
CspA_3xF_C_NcoI	*CCATGG*TTCTTAGATTTGAATCATTGATTTTAAC
CspA_3xF_D_BamHI	*GGATCC*GAATTAACAACAAATATAATGTATAAAAC
Construction of pMAD plasmid for chromosomic 3xFLAG-labeling of the *gdpP* gene	
GdpP_3xF_A_EcoRI	*GAATTC*TGAATCAACAGTGATGTATGCAG
GdpP_3xF_B_BamHI	*GGATCC*TTACTATTTATCGTCGTCATCTTTGTAGTCGATATCATGATCTTTATA ATCACCGTCATGGTCTTTGTAGTCTAATTGTTCTGTAATTGCTTGTTGT
GdpP_3xF_C_BamHI	*GGATCC*AGTAGGAGTGAAGATGCATGAA
GdpP_3xF_D_NcoI	*CCATGG*GTACCTTCAACTTCTTTATCTAATTTAACA
GdpP_3xF_E	CTAATCCATTGTTGATATATATGGAACC
GdpP_3xF_F	GCGGCATTTGATTTTGCT
Construction of plasmid expressing the SigB^{3xFLAG} protein	
EcoRI_sigB_fw	*GAATTC*AAAGAGCAGGTGCGAAATA
SpeI_sigB_rev	*ACTAGT*TTGATGTGCTGCTTCTTGTA
SpeI_3xF_TT_pCN51	*ACTAGT*GACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGACTA CAAAGATGACGACGATAAATAAGGCGCGCCTATTCTAAATG
NarI_3xF_TT_pCN51	AAA*GGCGCCT*GTCACTTTGC
Construction of plasmid expressing CspA and its variants	
CspA_+1_BamHI	*GGATCC*AAGCAGATGATTATTCCATATTG
CspA_ter_EcoRI	*GAATTC*CTATTGCAAACAATGTTGGTGATATAA
CspA_D5UTR_BamHI	*GGATCC*TCTTGGAGGTTTCAGATTATGAAACAAGGTACAGTTAA
CspX_D3UTR_KpnI	GGTACCAAAAAAAGAGTTCTCAATTTGGTTGAGAAACTCATCGCGTTAAGTAT TCTATTTATCGTCGTCATCTT
Construction of plasmid expressing CspB and its variants	
CspB_+1_BamHI	*GGATCC*AAGCTCGTGAATTAAATTGTAGTGTA
CspB_ter_EcoRI	*GAATTC*ATCATCGTGGTATCGTCTCATATT
3xFcspB_B	TTATAATCACCGTCATGGTCTTTGTAGTCAACAGTTTGTACGTTAACTGC
3xFCspB_C	GACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGACTACAAAG ATGACGACGATAAATAATCTTACAACATAAAACGACTCATTAT
CspB_D5UTR_BamHI	*GGATCC*TCTTGGAGGTTTCAGATTATGAAACAAGGTACAGTTAA
CspX_D3UTR_KpnI	GGTACCAAAAAAAGAGTTCTCAATTTGGTTGAGAAACTCATCGCGTTAAGTAT TCTATTTATCGTCGTCATCTT

Continued in the following page

Table 3. Continued

Construction of plasmid expressing CspC and its variants	
CspC +1 BamHI	*GGATCC*AATAAAGAGCGTGAAGAAAAATGTG
3xFcspC_B	ACCGTCATGGTCTTTGTAGTCCATTTTAACTACGTTTGCAGCTT
3xFcspC_C	GACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGACTACAAAG ATGACGACGATAAATAATTTTAACTTATTCAAACAGT
CspC ter KpnI	*GGTACC*CCATTATATATTGTTAAATCTCCAAC
CspC_D5UTR_BamHI	*GGATCC*TCTTGGAGGTTTCAGATTATGAATAACGGTACAGTTAA
CspX_D3UTR_KpnI	*GGTACC*AAAAAAAGAGTTCTCAATTTGGTTGAGAAACTCATCGCGTTAAGTAT TCTATTTATCGTCGTCATCTT
Synthesis of the *cspC mRNA*	
cspA+1_T7prom	**TAATACGACTCACTATAGGG**AATAAAGAGCGTGAAGAAAAATG
cspA_term	AAAAAATAACGGCTAGTGCTTTT
Synthesis of the *cspA* 5'UTR	
cspA+1_T7prom	**TAATACGACTCACTATAGGG**AAGCAGATGATTATTCCATATTG
rev_cspA_5UTR	AATCTGAAACCTCCAAGACTAAAA
ssDNA oligonucleotide gel shift assays	
f-A1	AAGCAGATGATTATTCCATATTGCAAAGAATATTTGAGTGAAATTTGCATGTA TTTTTTGTA
f-A2	TTTTTTGTAATATGCGAATAAGCATATTGAATGAATTTTAGTCTTGGAGGTTTC AGATTATG
f-A3	CTTGGAGGTTTCAGATTATGAAACAAGGTACAGTTAAATGGTTTAACGCTGAA AAAGGAT
f-A1-ΔT	AAGCAGATGATTATTCCATATTGCAAAGAATATTTGAGTGAAATTTGCATGTA
f-A2-ΔT	GTAATATGCGAATAAGCATATTGAATGAATTTTAGTCTTGGAGGTTTCAGATT ATG
f-C1	AATAAAGAGCGTGAAGAAAAATGTGAGTTATTTATATAGAATATTCTCCTTTTC ATTTAT
f-C2	CCTTTTCATTTATGAATTTGTTACAAAATATTTAGTGCAAAAGCACGACGGAG GTATTCAATATG
f-C3	CGACGGAGGTATTCAATATGAATAACGGTACAGTTAAATGGTTTAATGCAGA AAAAGGTT
f-C4	GAAAAAGGTTTTGGTTTCATCGAAAGAGAAGATGGTAGCGACGTATTCGTAC ACTTCTCA
f-C5	ACACTTCTCAGCAATCGCTGAAGATGGATACAAATCATTAGAAGAAGGCCAA AAAGTTGA
f-C6	AAAAAGTTGAATTCGACATCGTTGAAGGCGACCGTGGCGAGCAAGCTGCAA ACGTAGTTAAAA
f-C7	GTAGTTAAAATGTAATTTTAACTTATTCAAACAGTCCTTACTATAGGGCTGTTT TTTTAT

[a] Restriction enzymes sites, T7 promoter and 3xFLAG sequences included into the oligonucleotides are indicated in italic, bold and underlined format, respectively.

CHAPTER I:
S. aureus CSPs similarity is not linked to a functional redundancy

CHAPTER I:

S. aureus CSPs similarity is not linked to a functional redundancy.

As it was explained in the Introduction section, the number of genes encoding CSPs is variable in the different bacterial species. Unlike other protein families that include several homologs in same genome, such as the two-component systems and the GGDEF-domain (containing diguanylate cyclases), CSPs have a high protein sequence conservation among their members. This degree of similarity has long promoted speculations about the existence of redundant functions among CSPs.

Recent RIP-seq analysis developed in *Salmonella* revealed that CspC and CspE virulence-related RNA targets overlap almost entirely. Although single mutations of these CSPs did not pose any significant phenotypic changes, a double *cspC* and *cspE* gene mutation completely eliminated bacterial virulence (Michaux *et al.*, 2017). This phenotype could be complemented by overexpressing either one or the other CSP, confirming that both CspC and CspE have the same function. However, the other four CSPs in *Salmonella* bound different sets of targets, indicating independent

roles for each of them. This was in agreement with the idea that most CSPs display a particular function (Xia *et al.*, 2001; Michaux *et al.*, 2017). In this chapter, we analysed if *S. aureus* paralogs CspA, CspB and CspC, with a protein identity higher than 70 % (Fig. 1), could play the same role.

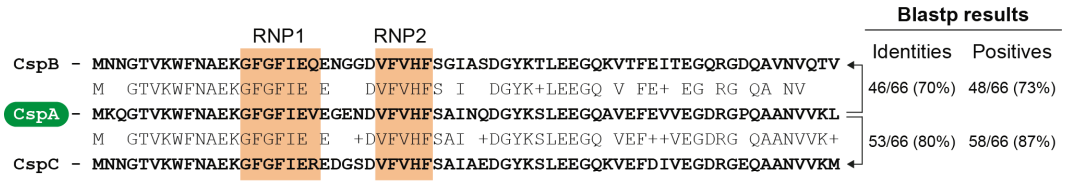

Figure 1. Comparison of *S. aureus* cold shock protein sequences. Blastp results of CspA vs CspB and CspA vs CspC alignments are shown. RNP1 and RNP2 motifs are highlighted in orange.

CspB and CspC paralogs do not complement CspA function when expressed from a heterologous promoter.

As previously described, CspA is required for an efficient production of staphyloxanthin (STX) (Katzif *et al.*, 2005). This yellow pigment functions as an antioxidant, which is relevant during infection as it offers protection against the innate immune response. The levels of STX can be easily measured by spectrophotometry after pigment extraction (Chia-I Liu *et al.*, 2008),(Lan *et al.*, 2010),(George Y Liu *et al.*, 2005). Taking advantage of this phenotype, we analysed if CspB or CspC could complement the STX production in a mutant strain, lacking the *cspA* gene. Hence, to control the expression of the CSP paralogs, we constructed the pCspA, pCspB and pCspC plasmids using pCN51 as a backbone (Charpentier *et al.*, 2004). Genes expressed from the pCN51 plasmid rely on the heterologous P*cad-CadC* module, which is inducible by cadmium, for their transcription. (Endo and Silver, 1995; Charpentier *et al.*, 2004). The Pcad promoter would avoid putative transcriptional regulatory mechanisms that may naturally occur to the chromosomal gene copies.

Each of the *csp* genes was expressed from the corresponding full-length mRNA, including the native 5' and 3'UTRs, for which transcripts were mapped using the transcriptome information from previous works (Lasa *et*

al., 2011; Koch *et al.*, 2014). As Figure 2 shows, the three *csp* genes contain long UTR regions.

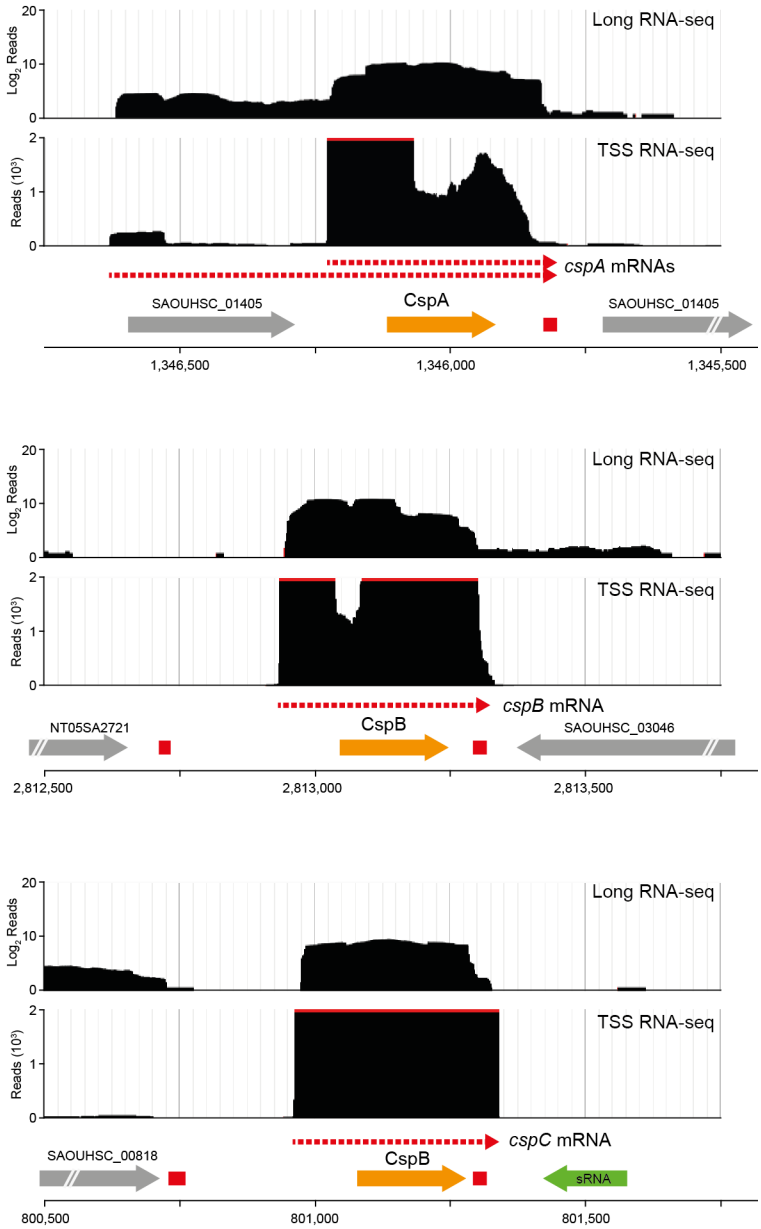

Figure 2. mRNAs of *S. aureus* CSPs contain long UTRs. Transcriptomic maps from a previous study showing the mRNA configuration of *S. aureus cspA*, *cspB* and *cspC* genes (Lasa *et al.*, 2011). In all cases the 5' and 3' UTRs cover a significant percentage of the mRNA length.

Plasmids were afterwards electroporated into the Δ*cspA* strain and STX extractions from pellets of bacteria grown in MH, supplemented with 1.5 mM of $CdCl_2$, performed. The quantification of the pigment production revealed that only Δ*cspA* carrying pCspA was able to reach the wild type (WT) levels. The complemented strains carrying pCspB and pCspC showed similar pigment levels to those found in the Δ*cspA* strain (Fig. 3). This result initially suggested that each of the *S. aureus* CSPs might have different functions.

Specific post-transcriptional regulation drives differential CSP expression.

It is well known that protein expression is influenced by transcriptional and post-transcriptional regulation. Therefore, the expression of CSPs from the same heterologous promoter did not guarantee the same protein levels in the complemented Δ*cspA* strains. In other words, we could not exclude that the *cspB* and *cspC* mRNAs were subjected to post-transcriptional regulatory mechanisms that could differentially affect their expression. Inadequate CSP amounts could lead to failure in complementing the phenotype.

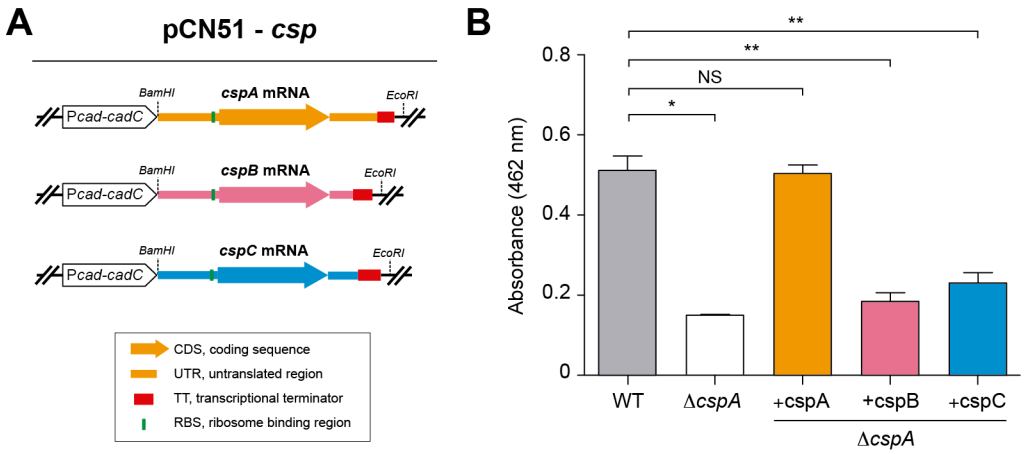

Figure 3. CspA paralogs (CspB and CspC) do not complement the staphyloxanthin (STX) production in a *cspA* deficient strain. A. Schematic representation of the pCN51 plasmid constructs harbouring the different *csp* mRNA sequences that were electroporated into the Δ*cspA* strain. Restriction enzymes used for cloning are indicated. **B.** Quantification of STX production in Δ*cspA* complemented with the pCspA, pCspB and pCspC plasmids. The STX pigment was extracted with ethanol from bacterial pellets after growth in MH for 16 hs at 37ºC. The WT and Δ*cspA* strains harbouring the pCN51 plasmid were used as positive and negative controls, respectively. Quantification was performed by measuring the absorbance of ethanol solutions at 462 nm. The plot shows the mean of absorbance from three independent pigment extractions and the standard deviation from Student's *t*-test, using the WT value as a reference. The asterisk indicates the statistical significance of the corresponding mean differences.

In order to correlate STX production with CSPs protein levels, we repeated the complementation of the Δ*cspA* strain using pCN51 plasmids carrying 3xFLAG-tagged CSPs (Fig. 4A). Quantification of STX levels from these tagged strains revealed similar results to the untagged ones (Fig. 4B). On the one hand, $CspA^{3xFLAG}$ was able to restore the pigment production, indicating that 3xFLAG did not affect the protein function. On the other hand, strains complemented with $CspB^{3xFLAG}$ and $CspC^{3xFLAG}$ showed a similar amount of pigment to that of Δ*cspA* (Fig. 4B). Interestingly, Western blot results revealed that $CspB^{3xFLAG}$ and $CspC^{3xFLAG}$ did not reach $CspA^{3xFLAG}$ protein levels, despite being expressed from the same *Pcad-CadC* module (Fig. 4C). Although CspC had the closest identity to CspA, it was also the least expressed, indicating the presence of different post-transcriptional regulatory mechanisms acting on each of the *csp* mRNAs. While this finding proves interesting, the strategy prevents us from addressing *S. aureus* CSPs putative functional redundancy.

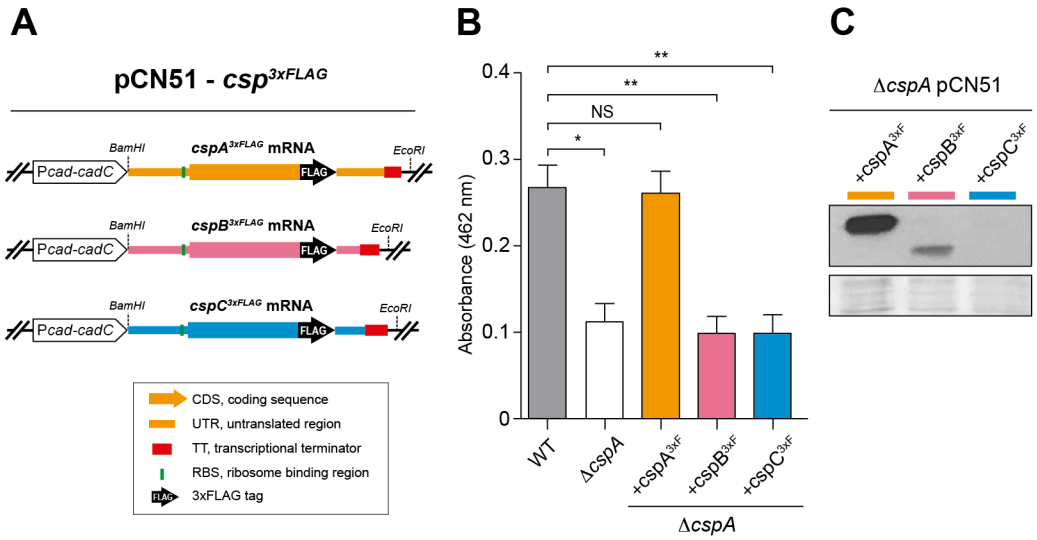

Figure 4. Differential post-transcriptional regulation of the CSP paralogs avoids complementation of the CspA phenotype. A. Schematic representation of the pCN51 plasmid constructs harbouring tagged CSP mRNAs. 3xFLAG was introduced at the C-terminus of the CDS. These plasmids were used to transform the Δ*cspA* strain. **B.** Quantification of STX production in Δ*cspA* complemented with the pCspA^{3xFLAG}, pCspB^{3xFLAG} and pCspC^{3xFLAG} plasmids. The WT and Δ*cspA* strains harbouring the pCN51 plasmid were used as positive and negative controls, respectively. The STX extraction was performed, quantified and plotted as indicated in Figure 3. **C.** Western blot showing the expression of the 3xFLAG-tagged CspA, CspB and CspC proteins in the Δ*cspA* strain. Total protein extractions were performed at the same time point and conditions described in B. Samples were run into 12% polyacrylamide gels and transferred to Nitrocellulose membranes. Tagged proteins were developed using peroxidase conjugated anti-FLAG antibodies and a bioluminescence kit. Coomassie stained gel portions are shown as loading controls.

The lack of complementation of a *cspA* deficient strain with CspB and CspC is independent of their protein levels.

In an attempt to avoid the putative post-transcriptional mechanisms affecting CspB and CspC protein expression, we decided to genetically modify *csp* mRNAs by eliminating UTRs. It has been previously described that both 5' and 3'UTRs contain post-transcriptional regulatory elements that have an impact on the protein levels (Waters and Storz, 2009; Ruiz de Los Mozos *et al.*, 2013; Ren *et al.*, 2017). We reasoned that elimination of these untranslated regions could help improving CSPs concentration inside the cell. Therefore, we constructed a new set of pCN51 plasmids expressing 3xFLAG-tagged *csp* mRNAs that lacked both UTRs while presenting the RBS and transcriptional terminator of the *cspA* mRNA (pCspA^{3xFLAG}ΔUTRs, pCspB^{3xFLAG}ΔUTRs and pCspC^{3xFLAG}ΔUTRs). These two common elements for all the constructs were included in an attempt to avoid differences in translation and termination of transcription (Fig. 5A). Surprisingly, elimination of the UTRs produced a significant reduction in the production of all CSPs (Fig. 5B). On the one hand, this result questioned the hypothesis of the presence of repressive regulatory mechanisms affecting the concentration of CspB and CspC through the UTRs. On the other hand, it supported the existence of a putative activator mechanism that improved CspA expression.

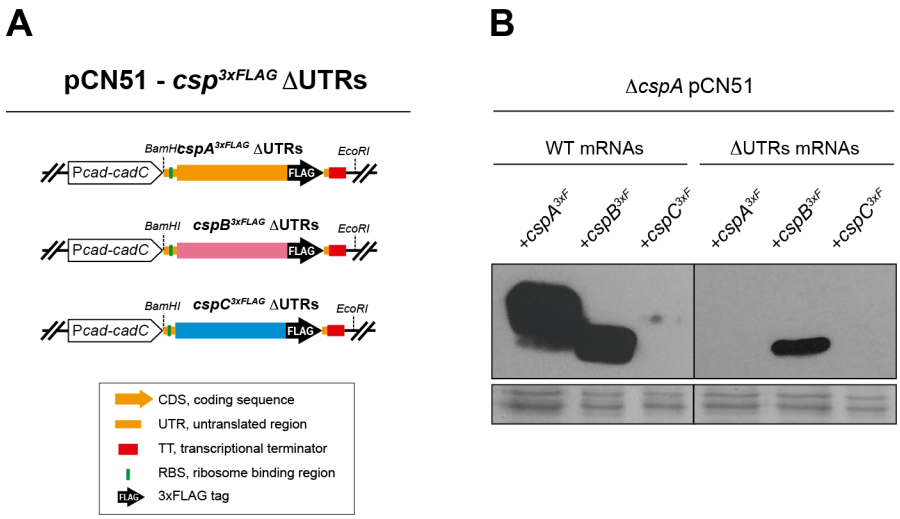

Figure 5. Elimination of the UTRs from the *csp* mRNAs does not increase their protein expression. A. Schematic representation of the pCN51 plasmid constructs harbouring *csp* mRNAs lacking the 5' and 3'UTRs. Only the RBS and TT from *cspA* mRNA were introduced as non-coding elements. These plasmids were used to transform the Δ*cspA* strain. **B.** Western blot showing the expression of the 3xFLAG-tagged CspA, CspB and CspC proteins in Δ*cspA*. Proteins were extracted, run, transferred and developed as described in Figure 4. Coomassie stained gel portions are shown as loading controls.

Since removing the UTRs from CSPs did not cause the expected effect, we decided to make mutations in the *cspA* the gene for it to codify CspB and CspC. This was achieved by performing the fewest nucleotide substitutions needed to transform $CspA^{3xFLAG}$ CDS into $CspB^{3xFLAG}$ and $CspC^{3xFLAG}$ CDSs, respectively (Table 4 and 5).

These quimeric mRNAs, which preserved the *cspA* 5' and 3'UTRs, were synthesized *in vitro* by GeneArt. Subsequently, we subcloned them into pCN51 and the resulting plasmids (pCspA_orfB^{3xFLAG} and pCspA_orfC^{3xFLAG}) were used to transform the Δ*cspA* strain (Fig. 6A).

Table 4. Codon mutations in the *cspA* mRNA to express the $CspB^{3xFLAG}$ protein

Position (aa)	CspA codon	Codon mutation[1]	Amino acid change
2	AAA	AA**T**	K2N
3	CAA	**AAT**	Q3N
20	GTT	**CAA**	V20Q
22	GGA	**AAT**	G22N
23	GAA	G**GC**	E23G
24	AAT	GG**T**	N24G
32	GCA	G**G**A	A32G
34	AAC	**GC**C	N34A
35	CAA	**TC**A	Q35S
40	TCA	**A**CA	S40T
46	GCT	**AAA**	A46K
48	GAG	**AC**G	E48T
51	GTA	**A**TA	V51I
52	GTT	**AC**T	V52T
55	GAC	**CA**A	D55Q
58	CCA	**GAT**	P58D
61	GCA	G**T**A	A61V
64	GTT	**CAA**	V64Q
65	AAA	A**C**A	K65T
66	CTA	**G**TA	L66V

[1]Nucleotides changes introduced in the corresponding codons are indicated in bold letters

Table 5. Codon mutations in the *cspA* mRNA to express the CspC^{3xFLAG} protein

Position (aa)	CspA codon	Codon mutation[1]	Amino acid change
2	AAA	AA**T**	K2N
3	CAA	**A**A**C**	Q3N
20	GTT	**AGA**	V20R
22	GGA	G**AT**	G22D
23	GAA	G**GT**	E23G
24	AAT	A**G**T	N24S
34	AAC	**GCT**	N34A
35	CAA	**G**AA	Q35E
46	GCT	**AAA**	A46K
50	GAA	GA**T**	E50D
51	GTA	**A**TA	V51I
58	CCA	**GA**A	P58E
66	CTA	**A**T**G**	L66M

[1]Nucleotides changes introduced in the corresponding codons are indicated in bold letters

Western blots showed that CspA^{3xFLAG}, CspB^{3xFLAG} and CspC^{3xFLAG} proteins were produced at similar rate (Fig. 6B). This result confirmed that *cspB* and *cspC* mRNAs were subjected to unknown post-transcriptional mechanisms and excluded post-translational processes as responsible for modifying protein levels. Nevertheless, quantification of the STX levels revealed that the production of the yellow pigment could not be restored in the Δ*cspA* strain by complementation with CspB and CspC, indicating the existence of a specific role for each *S. aureus* CSP (Fig. 6C).

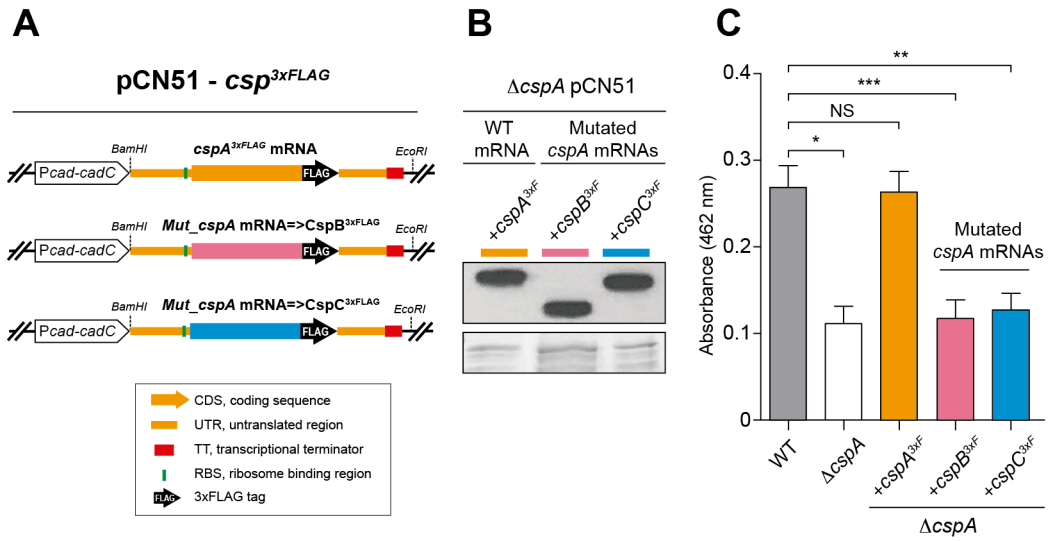

Figure 6. Increased CspB and CspC levels do not complement staphyloxanthin (STX) production in a *cspA* deficient strain. A. Schematic representation of the pCN51 plasmid constructs harbouring modified *cspA* mRNAs, which were generated to encode CspB or CspC labelled with the 3xFLAG at the C-terminus. These plasmids were introduced in the Δ*cspA* strain. **B.** Western blot showing the expression of 3xFLAG-tagged CspA, CspB and CspC proteins in Δ*cspA*. Proteins were extracted, run, transferred and developed as described in Figure 4. **C.** Quantification of STX production in Δ*cspA* complemented with pCspA^{3xFLAG}, pCspA_orfB^{3xFLAG} and pCspA_orfC^{3xFLAG} plasmids. WT and Δ*cspA* harbouring the pCN51 plasmid were used as positive and negative controls, respectively. The STX pigment was extracted, quantified and plotted as indicated in Figure 3.

The RNP-connecting loop is not responsible for CSP specificity

It is widely accepted that the RNA binding capacity of CSPs resides in the RNP1 and RNP2 motifs. These amino acids are highly conserved in most organisms, including *S. aureus* (Fig. 1). However, those that are responsible for target specificity remain unknown. Some studies suggest that the charge of the protein surface, the different residues surrounding RNP1 and RNP2 motifs or the non-conserved amino acidic loop connecting RNP1 and RNP2 could drive target specificity (Kenan *et al.*, 1991; Yamanaka *et al.*, 1998; Phadtare and Inouye, 1999).

In an effort to understand which region of *S. aureus* CspA was responsible for target specificity, we first considered the connecting loop between RNP1 and RNP2. The amino acid differences in this region were relatively high (4 out of 6 amino acidic changes) compared to the overall protein sequence variation. To test if the functionality of CspA was altered when carrying the RNPs-connecting loop sequences of CspB and CspC, we performed the following amino acid exchanges: V20Q, G22N, E23G and N24G, generating the CspA_MLB protein, and V20R, G22D, E23G and N24S, generating the CspA_MLC variant (Fig 7A). After synthetizing such CspA chimeras *in vitro,* we cloned them in pCN51, generating plasmids pCspA_MLB and pCspA_MLC, and used them to transform the Δ*cspA* strain. Surprisingly, both chimeric proteins restored the wild type

phenotype, with CspA_MLC showing even higher STX levels than the WT strain (Fig. 7B). This result demonstrated that the amino acid variations present in the RNPs-connecting loop are not responsible for the *in vivo* CSP specificity but may affect binding affinity in *S. aureus*.

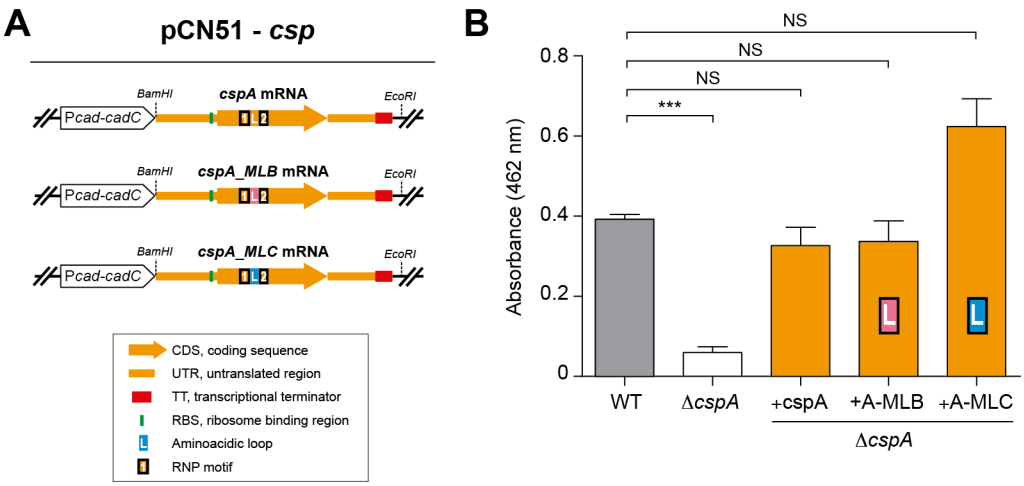

Figure 7. The RNPs-connecting loop is not responsible for CspA specificity. A. Schematic representation of the pCN51 plasmid constructs harbouring *cspA* mRNAs, which were modified to encode the CspA protein including the RNPs-connecting loop of CspB or CspC. These plasmids were used to electroporate the Δ*cspA* strain. **B.** Quantification of staphyloxanthin (STX) production in the Δ*cspA* strains complemented with pCN51-*cspA*, *_cspA_MLB* and *_cspA_MLC* plasmids. WT and Δ*cspA* harbouring the pCN51 plasmid were used as positive and negative controls, respectively. The STX pigment was extracted, quantified and plotted as indicated in Figure 3.

CspA target specificity may be encoded in the carboxi-half of the protein

The previous results suggested that the target specificity should be encoded outside of the RNA binding region. Considering that RNP1 and RNP2 motifs included into β2 and β3 strands are located in the first half of the protein, we wondered if the second half of the protein could play a role recognizing the targets. Thus, we constructed two additional chimeric *cspA* mRNAs expressing a CSP^{3xFLAG} that comprised the N-terminal half of CspA (which included the RNP1 and RNP2 motifs) and the C-terminal half of CspB or CspC ($pCspAB^{3xFLAG}$ and $pCspAC^{3xFLAG}$) following the codon substitution criteria applied above (Tables 6 and 7) (Fig 8A).

Table 6. Codon mutations included into *cspA* mRNA to express the chimeric $CspAB^{3xFLAG}$ protein

Position (aa)	CspA codon	Codon mutation[1]	Amino acid change
32	GCA	G**G**A	A32G
34	AAC	**GC**C	N34A
35	CAA	**TC**A	Q35S
40	TCA	**A**CA	S40T
46	GCT	**AAA**	A46K
48	GAG	**AC**G	E48T
51	GTA	**A**TA	V51I
52	GTT	**AC**T	V52T
55	GAC	**C**A**A**	D55Q
58	CCA	**GAT**	P58D
61	GCA	G**T**A	A61V
64	GTT	**CAA**	V64Q
65	AAA	A**C**A	K65T
66	CTA	**G**TA	L66V

[1]Nucleotides changes introduced in the corresponding codons are indicated in bold letters

Table 7. Codon mutations included into *cspA* mRNA to express the chimeric CspAC^{3xFLAG} protein

Position (aa)	CspA codon	Codon mutation[1]	Amino acid change
34	AAC	**GCT**	N34A
35	CAA	**G**AA	Q35E
46	GCT	**AAA**	A46K
50	GAA	GA**T**	E50D
51	GTA	**A**TA	V51I
58	CCA	**GA**A	P58E
66	CTA	**AT**G	L66M

[1]Nucleotides changes introduced in the corresponding codons are indicated in bold letters

Although the protein levels of these chimeras were comparable to the original CspA^{3xFLAG} protein, the analysis of STX production of the Δ*cspA* strains harbouring these plasmids revealed that the pigment levels could not be restored (Fig. 8B and 8C). This indicated that the few amino acid differences in the C-half of the CSPs may be responsible for CSP specificity *in vivo*, at least in *S. aureus*.

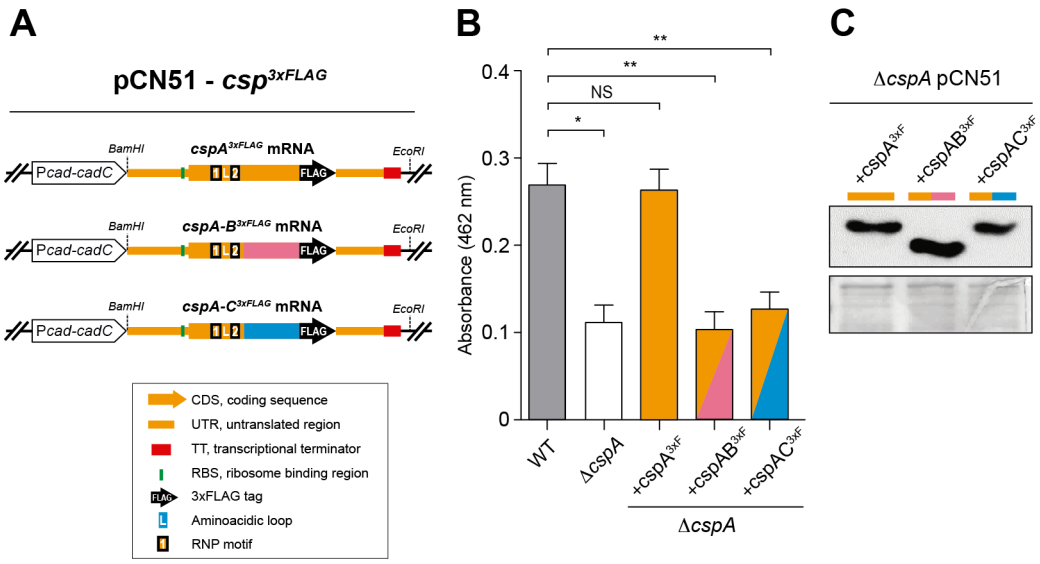

Figure 8. CspA carboxi-half amino acids are required to complement staphyloxanthin (STX) production in a *cspA* deficient strain. A. Schematic representation of the pCN51 plasmid constructs harbouring *cspA* mRNAs that were modified to encode chimeric CspAB and CspAC proteins labelled with the 3xFLAG at the C-terminus of the CDS. These plasmids were introduced in the Δ*cspA* strain. **B.** Quantification of STX production in the Δ*cspA* strains complemented with pCspA^{3xFLAG}, pCspAB^{3xFLAG} and pCspAC^{3xFLAG} plasmids. The WT and Δ*cspA* strains harbouring the pCN51 plasmid were used as positive and negative controls, respectively. The STX pigment was extracted as indicated in Figure 3. **C.** Western blot showing the expression of 3xFLAG-tagged CspA, CspAB and CspAC proteins in the Δ*cspA* strain. Proteins were extracted, run, transferred and developed as described in Figure 4.

CHAPTER II:
The regulon of the staphylococcal RNA chaperone CspA

CHAPTER II:

The regulon of the staphylococcal RNA chaperone CspA.

CspB and CspC inability to restore the STX phenotype in a Δ*cspA* strain, as described in Chapter I, suggested that each of the CSPs has a specific function in *S. aureus*. Since these CSPs are generally accepted as RNA binding proteins, one could speculate that *S. aureus* variants have unique targets and, overall, different effects on the biology of the bacteria. In this Chapter, we focused on determining the regulon of *S. aureus* CspA using genome-wide approaches.

S. aureus CspA is a global modulator of gene expression

In order to evaluate the biological role of CspA in *S. aureus*, we performed comparative label-free LC-MS-based proteomics of the WT and Δ*cspA* strains. For this purpose, we extracted total protein from both strains at mid-exponential phase in biological triplicates. After trypsin treatment, the digested peptides were analysed by LC-MS/MS and identified by ProteinPilot Software (using the *S. aureus* NCTC 8325 protein database). The abundance of the identified proteins in the WT and Δ*cspA* strains was

determined by Progenesis IQ software, which quantified a total of 1,206 proteins (a coverage of 43.6% of the *S. aureus* NCTC 8325 proteome). We only considered non-conflicting peptides with a confidence identification score higher than 3, an error mass (ppm) lower than 20 and an m/z ratio between 450 and 800. Among them, 282 (80 up-regulated and 202 down-regulated) proteins were differentially expressed in the absence of CspA, with a fold-change ratio higher than 2 and a P-value lower than 0.05 (Fig. 9 and Annex I). Up- and down-regulated proteins were classified into 21 different functional categories, following the SEED database standards (http://pseed.theseed.org/) (Overbeek *et al.*, 2005). Proteins related to carbohydrate, nucleotide and protein metabolism, virulence and stress response were the most abundant (Fig. 10A). To verify if CspA had a bias for modulating expression of proteins belonging to a particular biological process, we performed GO terms enrichment analysis using the PANTHER Overrepresentation Test tool from the Gene Ontology Consortium (Fig. 10B) (http://pantherdb.org) (Mi *et al.*, 2017). This analysis revealed that proteins involved in processes such as carbohydrate and ribonucleotide biosynthesis (P-value = 8.81e-04 and 6.40e-05, respectively), small molecule metabolism (P-value = 6.50e-11) and oxidative-reduction (P-value = 1.17e-02), among others, were enriched in the down-regulated gene group. In contrast, proteins involved in *S. aureus* pathogenesis (P-value = 1.89e-03) and cytolysis (P-value = 1.41e-02), were enriched in the up-regulated group. The fact that clusters

of proteins belonging to specific biological processes were significantly affected in Δ*cspA* suggested that CspA might have a specialized function.

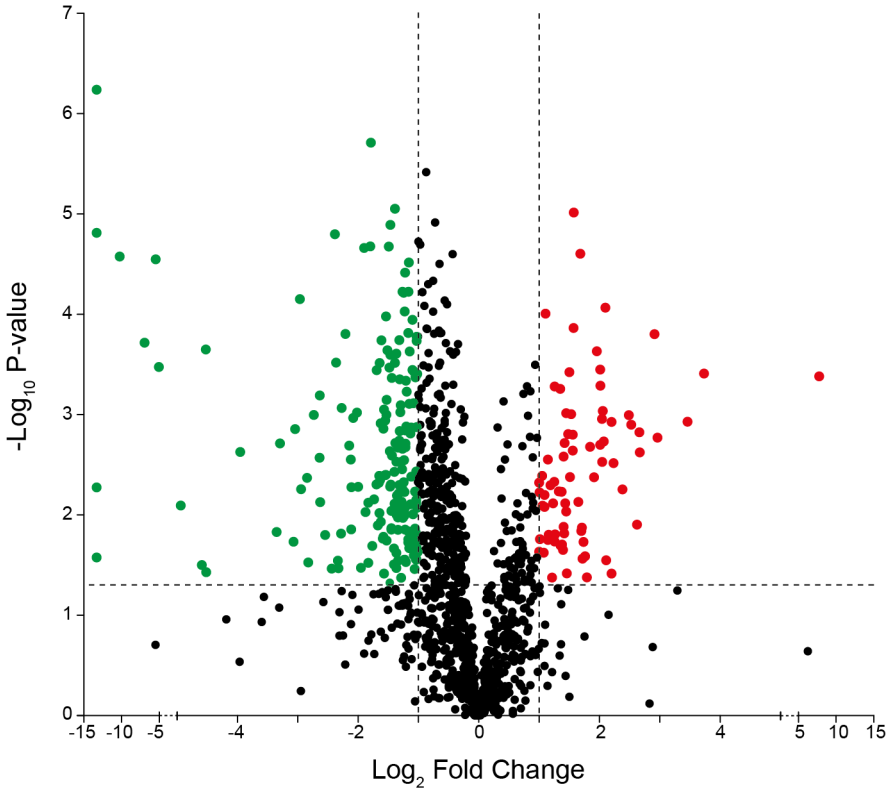

Figure 9. Volcano plot for differential gene expression determined by label-free proteomics in Δ*cspA* compared to the WT strain. Scattered points represent genes: the x-axis reflects the log ratio, log_2 fold change, between the Δ*cspA* and WT strains; whereas the y-axis indicates the negative log of the P-value. Red and green dots are the proteins that were significantly up- and down-regulated (considering a fold change > 2 and a P-value <0.05), respectively. Detailed information about these proteins is available in Annex I.

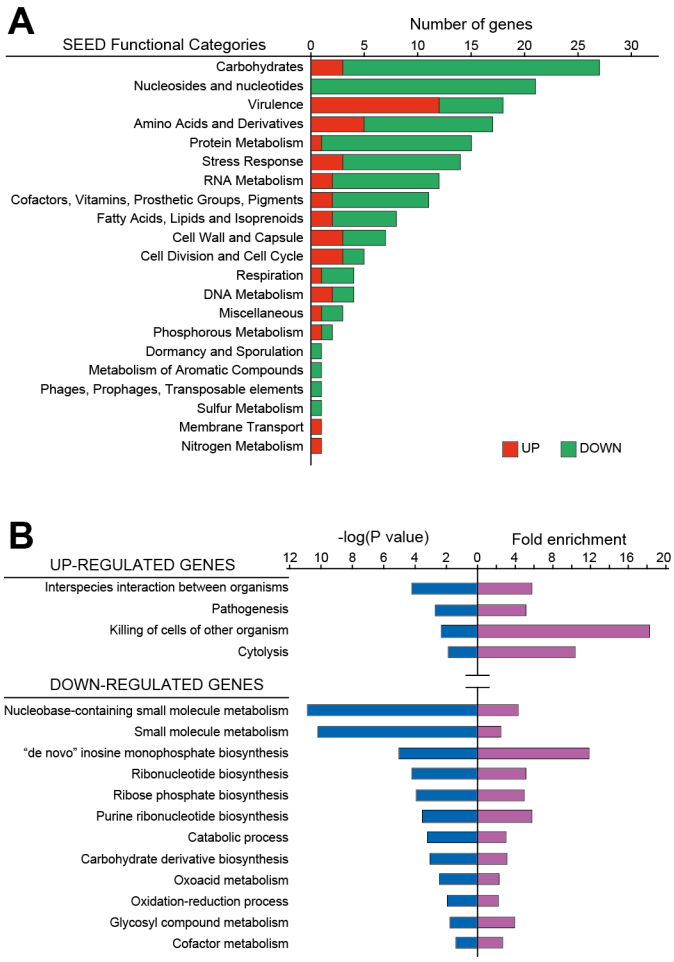

Figure 10. Functional classification and analysis of proteins affected by *cspA* deletion. A. The plot represents the number of detected up- and down-regulated proteins (red and green bars, respectively) that could be classified into different functional SEED categories (http://pseed.theseed.org/) (Overbeek *et al.*, 2005). **B.** Gene Ontology Enrichment Analysis performed with the PANTHER Overrepresentation Test tool from the Gene Ontology Consortium (http://pantherdb.org) (Mi *et al.*, 2017). P-value (represented as negative log) and fold enrichment of overrepresented functional categories (P-value <0.05) are plotted. To simplify the plot, redundant categories were eliminated.

CspA is required for modulation of stress-associated phenotypes

The previous analysis indicated that CspA is required for the proper expression of several biologically relevant genes in *S. aureus*. Therefore, we found it reasonable for some *S. aureus* phenotypes to be affected. There were two evident phenotypic changes that arose upon deletion of *cspA*: an increased bacterial aggregation and, as explained in Chapter I, reduced pigmentation. Concerning aggregation, the main exopolysaccharide involved in *S. aureus* biofilm formation is poly-N-acetyl-glucosamine (PIA-PNAG), whose synthesis depends on the expression of the *icaADBC* operon (Cramton *et al.*, 1999; Maira-Litran *et al.*, 2005). Quantification by dot-blot of PIA-PNAG indicated that such increased bacterial aggregation (Fig. 11A) could be attributed to a rise in PNAG production in Δ*cspA* (Fig. 11B), which correlated with augmented IcaB protein levels (FC = 7.5; P-value 0.0002) (Annex I). The reduced pigmentation can be explained by a decrease in the expression of the *crtOPQMN* operon, responsible for STX biosynthesis and controlled by the alternative sigma factor SigB (Pelz *et al.*, 2005). In agreement with previous results (Katzif *et al.*, 2005), the Δ*cspA* strain expressed lower levels of SigB (FC = -3.1; P-value 0.0003; Annex I). Reduction of SigB also correlated with lower levels of other proteins controlled by SigB, including the Asp23 (FC = -7.8, P-value = 0.00007; Annex I and Fig 11C),

which, alongside STX production, is typically used as an indicator of SigB activity (Kuroda *et al.*, 1995).

In addition to STX pigment, further proteins involved in oxidative-reduction processes, such as superoxide dismutase (FC = -2.0, P-value = 0.001), ferritin-like antioxidant (FC = -2.8, P-value = 0.00002) and glutathione hydrolase (FC = -2.0, P-value = 0.001), were significantly enriched among the down-regulated genes (Fig. 10B). Likewise, regulatory proteins such as MgrA, SarZ, and Fur involved in oxidative stress adaptation, as previously described, were less abundant in the Δ*cspA* strain (Annex I) (Gaupp *et al.*, 2012). Thus, to verify if the absence of CspA might affect oxidative-stress adaptation of *S. aureus*, the WT and Δ*cspA* strains were grown until late stationary phase and challenged with H_2O_2 for 1 h. We found that the number of viable bacteria recovered from Δ*cspA* was around one log lower than that from the WT strain, indicating that *S. aureus* CspA improves bacterial survival to oxidative stress (Fig. 11D). All the phenotypes associated to Δ*cspA* that we analysed were complemented by heterologous expression of the *cspA* gene, confirming the role of CspA in modulating their levels (Fig. 11). Altogether, these results illustrate the importance of CspA as a key player in ensuring the proper expression of important stress-related genes in *S. aureus*.

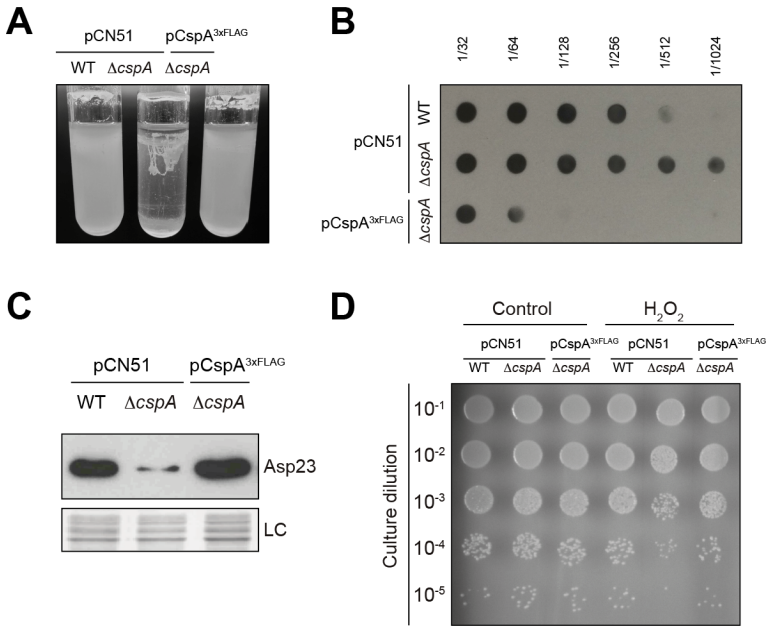

Figure 11. Phenotypic comparison of the wild type, Δ*cspA* and Δ*cspA* pCspA^{3xFLAG} strains. A. Bacterial aggregation phenotypes after incubation for 24 h at 250 rpm and 37ºC in glass tubes containing 5 ml of TSBglu. **B.** Dot-blot of PIA-PNAG exopolysaccharide for each of the strains in the previously mentioned conditions. Serial dilutions (1/2) of the samples were spotted onto nitrocellulose membranes and PIA-PNAG was developed with specific anti-PIA-PNAG antibodies. **C.** Asp23 levels. Total protein extraction was performed at mid-exponential phase after growth at 37ºC and 200 rpm. The Western blot was developed using peroxidase conjugated anti-FLAG antibodies. A Coomassie stained gel portion is shown as loading control (LC). **D.** Hydrogen peroxidase susceptibility assay. Bacterial plate growth from several ten-fold serial dilutions after treatment with a fixed concentration of H_2O_2. Bacteria were grown until late stationary phase at 37ºC and 200 rpm, diluted to $5x10^7$ cfu ml-1 and challenged with 0.09 % final concentration of H_2O_2 for 1 h in the same conditions. Non-treated bacteria were included as a control.

The *in vivo* targetome map of the *S. aureus* RNA chaperone CspA

In order to identify which of the regulated proteins found in the proteomic analysis were direct targets of CspA *in vivo*, we performed RIP-chip analysis (Jain *et al.*, 2010). Since the three copies of CSPs contained in the *S. aureus* genome are highly similar, the possibility to obtain polyclonal antibodies, able to specifically differentiate CspA from CspB and CspC proteins, seemed unlikely. Thus, we labelled CspA with a 3xFLAG translational fusion (15981 *cspA*^{3xFLAG}). Prior to introducing the tag, we modelled *S. aureus* CspA structure using Swiss-Model workspace (http://swissmodel.expasy.org) (Arnold *et al.*, 2006). Both N- and C-terminal ends were exposed on the protein surface and opposite to the RNP1 and RNP2 RNA binding domains, seeming equally suitable for the introduction of the flag (Newkirk *et al.*, 1994; Schindelin *et al.*, 1994) (Fig. 12A). We chose to include it at the C-terminus and tested the functionality of CspA^{3xFLAG} by verifying that 15981 *cspA*^{3xFLAG} displayed a similar amount of STX compared to the WT strain (Fig. 12B). Western blots showed that CspA^{3xFLAG} was expressed at all tested points of the growth curve (Fig. 12C).

In order to purify the RNAs associated to CspA^{3xFLAG}, we included the WT strain (lacking the 3xFLAG sequence) and the 15981 GdpP^{3xFLAG} strain (carrying a flagged version of GdpP, a protein without RNA-binding

domains), as negative controls. We included the latter as an additional negative control to exclude unspecific transcripts that might be pulled down alongside the 3xFLAG-anti-FLAG-G-sepharose complex. Western blots confirmed the presence of the 3xFLAG tagged proteins, CspA^{3xFLAG} and GdpP^{3xFLAG} in the purified fractions (Fig. 12D).

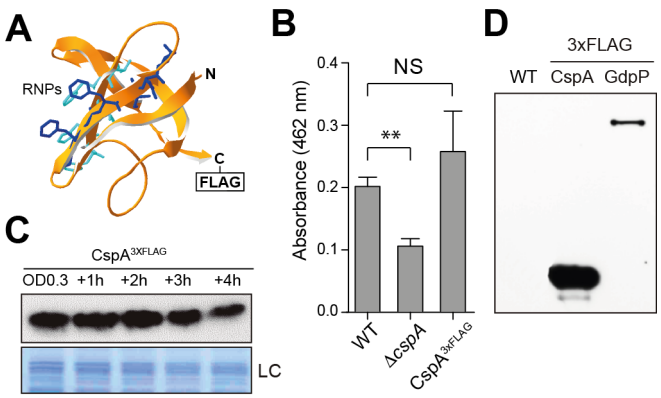

Figure 12. Tagging CspA protein does not affect its functionality and allows its co-immunoprecipitation. A. Putative *S. aureus* CspA protein structure, predicted by SWISS-MODEL (Arnold *et al*., 2006), using *B. subtilis* CspB as a template (PDB ID 2ES2). Residues corresponding to RNA-binding domains RNP1 and RNP2 are showed in blue and cyan respectively. **B.** Quantification of staphyloxanthin (STX) production in the WT, Δ*cspA* and CspA^{3xFLAG} strains. The STX pigment was extracted, quantified and plotted as indicated in Figure 3. **C.** Chromosomal CspA^{3xFLAG} expression profile along the growth curve. Protein samples were extracted when bacteria reached OD_{600nm} 0.3 and +1, +2, +3 and +4 h later, after growth at 37ºC and 200 rpm. The Western blot was developed using peroxidase conjugated anti-FLAG antibodies. A Coomassie stained gel portion is shown as loading control (LC). **D.** CspA^{3xFLAG} and GdpP^{3xFLAG} pull-down control. Western blot of the precipitated fractions showing the presence of flagged proteins after RIP was performed. The result was developed using peroxidase conjugated anti-FLAG antibodies and a bioluminescence kit.

We extracted RNAs bound to these purified proteins and identified them with the help of *S. aureus* custom tiling microarray chips, as previously described (Segura *et al.*, 2012). To visualize normalized CspA-binding signals we used the Integrated Genome Browser (IGB) (Freese *et al.*, 2016). The majority of the signals were only present in the $CspA^{3xFLAG}$ RIP-chip while just a few of them were also found in the negative controls (Fig. 13A). This indicated that $CspA^{3xFLAG}$ pull-down was specific. Strikingly, the pull down revealed signals in the form of peaks instead of full-length transcripts. This was probably due to RNA degradation occurring during RIP sample processing. Therefore, the peaks might represent RNA regions protected by CspA binding.

Next, we performed peak calling using two complementary methods, Thresholding and CisGenome (see Material and Methods), and identified 570 and 355 peaks, respectively. The difference in the number of peaks detected by both bioinformatics approaches can be explained by their distinct data processing procedures. Thresholding method considers all peaks above a certain level to be true while CisGenome tends to group contiguous peaks, representing them as only one. In this study, we only considered those regions that were commonly detected by both methods. Figure 13B, shows CspA-binding signals as broadly distributed across the whole staphylococcal genome. We integrated this results with the transcriptomic data, previously generated (Lasa *et al.*, 2011), and loaded

them into a public web server based on Jbrowse (Skinner *et al.*, 2009) (http://rnamaps.unavarra.es/).

To match the identified CspA-binding signals with their corresponding transcripts, we manually annotated the boundaries of each of the targeted RNAs based on our transcriptomic data (Lasa *et al.*, 2011; Ruiz de Los Mozos *et al.*, 2013). We observed that CspA-binding peaks included at least 213 transcripts of different RNA nature: monocystronic, polycystronic and non-coding transcripts, such as small RNAs or riboswitches (Annex II). After defining the summit position for each of the peaks on mRNAs, we found that most of the summit peaks (257) mapped onto CDSs and 59 onto untranslated regions (Fig. 13C). As Figure 13C shows, CspA-binding signals were homogeneously distributed along the targeted mRNAs, indicating no particular preference for a specific position within them. We also wondered if CspA recognised certain RNA sequences within these regions. Hence, we looked for a consensus RNA sequence domain using MEME suite Version 4.10.1 (Bailey *et al.*, 2009). We run the algorithm with the sequences covered by the CspA-binding peaks or with 60-nt sequences centred at the summit position. After testing different parameters, we did not find any clear RNA motifs that could be considered a consensus CspA binding region. Although an RNA binding sequence is still missing, these results revealed hundreds of putative RNA targets, depicting CspA as a potential global post-transcriptional modulator.

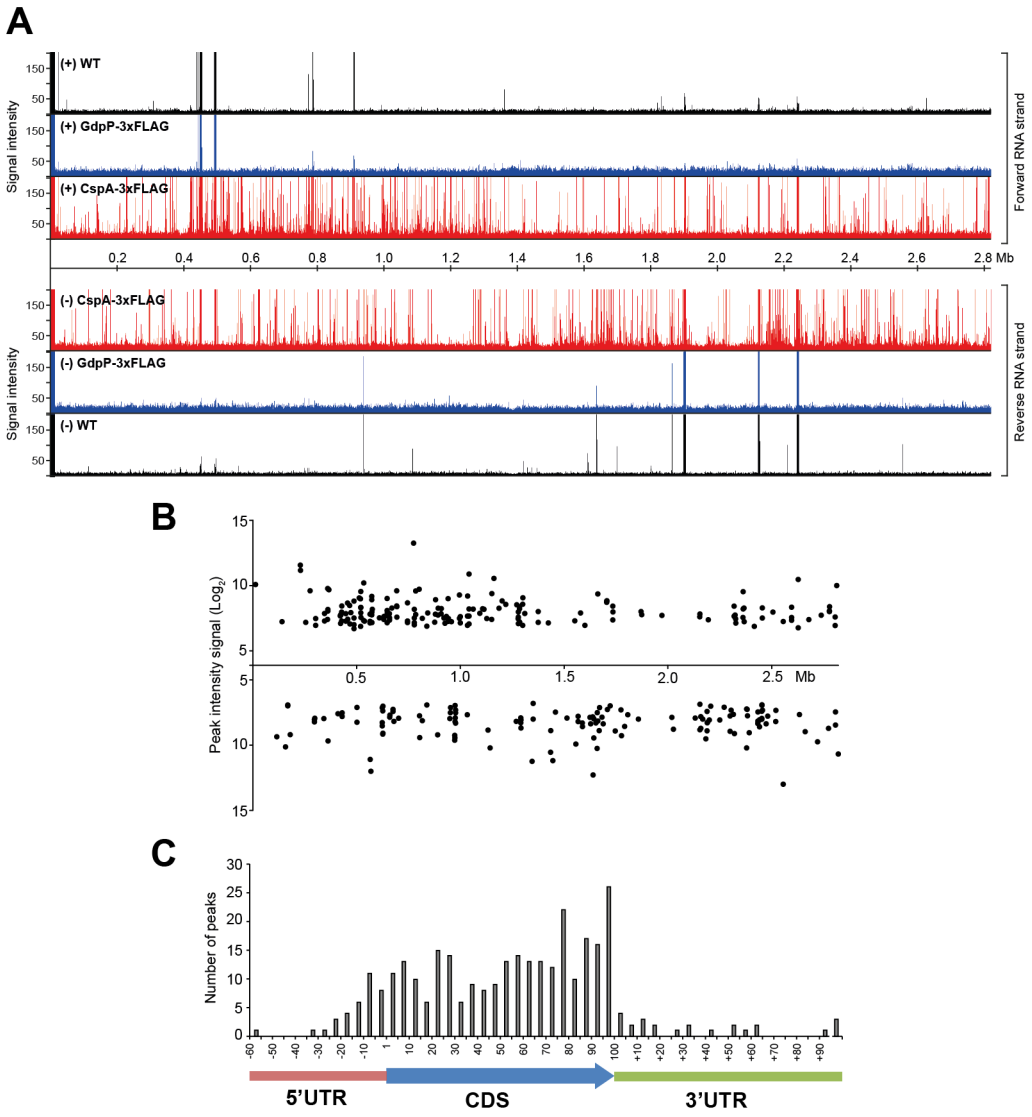

Figure 13. Targetome map of the RNA chaperone CspA in *S. aureus*. A. RIP-on-chip genomic maps. IGB plot showing the tiling array signals of the RIP-on-chip assays from strains $CspA^{3xFLAG}$ (red), $GdpP^{3xFLAG}$ (blue) and the untagged WT (black) along the forward and reverse genomic strands. *S. aureus* NCTC 8325 strain sequence was used as reference genome. These maps can be browsed at http://rnamaps.unavarra.es/. **B.** CspA binding peaks mapped across the *S. aureus* NCTC 8325 genome. Each dot represents the peak intensity signal for a specific genome position. **C.** Relative summit peak positions mapped onto an mRNA model. The length of CDSs encoded by transcripts targeted by CspA were normalized to 100 and the summit positions were mapped accordingly. The number of relative summit positions mapping onto each 1/20 fraction of a CDS or outside of it were plotted.

CspA binding can positively or negatively modulate the expression of its targets at the post-transcriptional level

Integration of the proteomic and the targetome data revealed 52 transcripts that were directly bound by CspA and had significantly different protein levels in Δ*cspA* (17 up-regulated and 35 down-regulated) (Annex III). Among them, 36 candidates were classified into SEED categories, with the stress response group being one of the most represented, including: cold shock protein CspC, the alternative RNA polymerase sigma factor SigB, serine-protein kinase RsbW and manganese superoxide dismutase SOD (Fig. 14 and Annex III). In order to validate the putative post-transcriptional effect on CspA targets due to CspA binding, we selected SigB and CspC as model examples for down- and up-regulated proteins, respectively (Fig. 15A). In addition, we also included CspA since the RIP-chip analysis revealed binding signals onto its own mRNA (Fig. 15A). To exclusively monitor protein expression at the post-transcriptional level, we tagged the selected genes with 3xFLAG and placed them under the control of heterologous promoters. The resulting plasmids (pSigB^{3xFLAG}, pCspC^{3xFLAG}, pCspA^{3xFLAG}) were used to transform the WT and Δ*cspA* strains and their protein expression analysed by Western blot at mid-exponential growth phase.

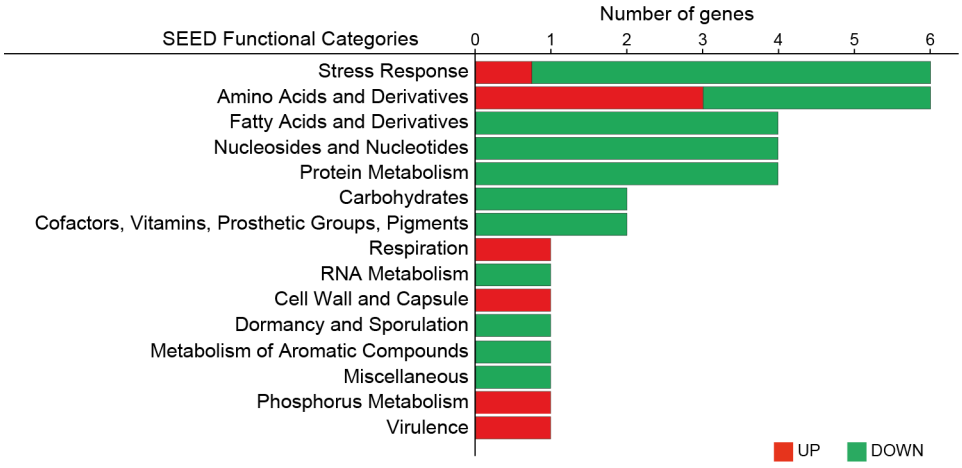

Figure 14. Targetome map of the RNA chaperone CspA in *S. aureus*. A. Functional classification of differentially expressed CspA targets. The plot represents the number of differentially expressed CspA targets that could be classified into different functional SEED categories (http://pseed.theseed.org/) (Overbeek *et al.*, 2005). The number of up-regulated and down regulated targets is shown as red and green bars, respectively. Details of the represented genes are included in Annex I and Annex III.

As shown in Figure 15B, $SigB^{3xFLAG}$ and $CspC^{3xFLAG}$ expression in the Δ*cspA* strain was lower and higher, respectively, compared to that in the WT strain. This was in agreement with the proteomic data (Annex I; SigB FC = -3.1, P-value = 0.0003 and CspC FC = 11.00, P-value 0.001), indicating that CspA can modulate both positively and negatively the expression of its targets. Like CspC, the expression of $CspA^{3xFLAG}$ was higher in the Δ*cspA* than in the WT strain, suggesting a possible post-transcriptional negative loop acting on the *cspA* mRNA (Fig. 15B). These findings proved relevant since the RNA chaperone activity of CSPs is expected to contribute to an improvement of translation efficiency (by melting RNA secondary structures that impair ribosome movement) and not the contrary.

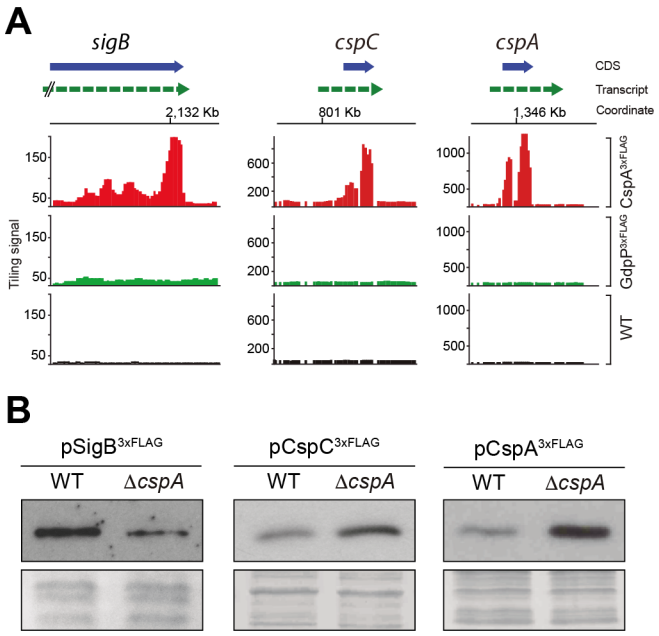

Figure 15. Post-transcriptional regulation of selected CspA targets. A. RIP-chip maps showing CspA-binding signals on the *sigB*, *cspC* and *cspA* loci. Normalized tiling signals of RIP-on-chip experiments for CspA3xFLAG, GdpP3xFLAG and WT are shown as red, green and black bars respectively. CDSs appear as blue box arrowsy and mRNAs are represented as dashed green arrows. **B.** Expression of 3xFLAG-tagged SigB, CspC and CspA proteins in the WT and Δ*cspA* strains. Total protein extraction was performed at mid-exponential phase after growth at 37ºC and 200 rpm. Samples were run into 12% polyacrylamide gels and transferred to Nitrocellulose membranes. Western blots were developed using peroxidase conjugated anti-FLAG antibodies and a bioluminescence kit. Coomassie stained gel portions are shown as loading controls (LC).

CHAPTER III:
CspA represses CspC translation

CHAPTER III:

CspA represses CspC translation.

Considering that some of CspA direct targets had increased protein levels in its absence (Chapter II and Annex III), we concluded that the CspA binding had a negative effect on the expression of such targets. This proved relevant, since it is commonly expected from CSPs chaperone activity to increase translation efficiency (by melting RNA secondary structures that impair ribosome movement). For this reason and to gain further knowledge on the capacity of CspA to inhibit protein expression at the post-transcriptional level, we chose the *cspC* transcript as a model.

CspA protein binds the 5'UTR of the *cspC* mRNA *in vitro*

Confirming that CspA effectively bound the *cspC* mRNA was our first priority, for which we decided to perform electrophoretic mobility shift assays (EMSAs). Since we needed purified recombinant CspA (rCspA), we first generated a plasmid expressing a GST-tagged CspA protein under the control of the *lacZ* promoter (pGEX-6P-2::cspA), which we subsequently used to transform *E. coli* BL21 cells (*E. coli* BL21 pGEX-6P-

2::cspA). We then grew *E. coli* BL21 pGEX-6P-2::cspA strain at 37ºC and 250 rpm until an OD_{600} of 0.5 was reached, added 0.4 mM IPTG and incubated them for an additional 5 h. Afterwards, we purified the GST-CspA fusion protein from clarified lysates, using a GSTrap column coupled to an AKTAprime plus chromatography system, and collected several fractions from the lysate after flowing through the column. We run a small portion of them in acrylamide gels. Figure 16 shows a Coomasie stained gel with the relative amount of total protein obtained before and after IPTG induction, and once treatment with Precision protease was complete (used to cleave the GST-CspA fusion protein and release rCspA). An enriched band corresponding to GST-CspA indicates an increase in the post-induction sample compared to the pre-induction sample. In the purified fractions, single discrete bands corresponding to rCspA reflect that the purification process was successful.

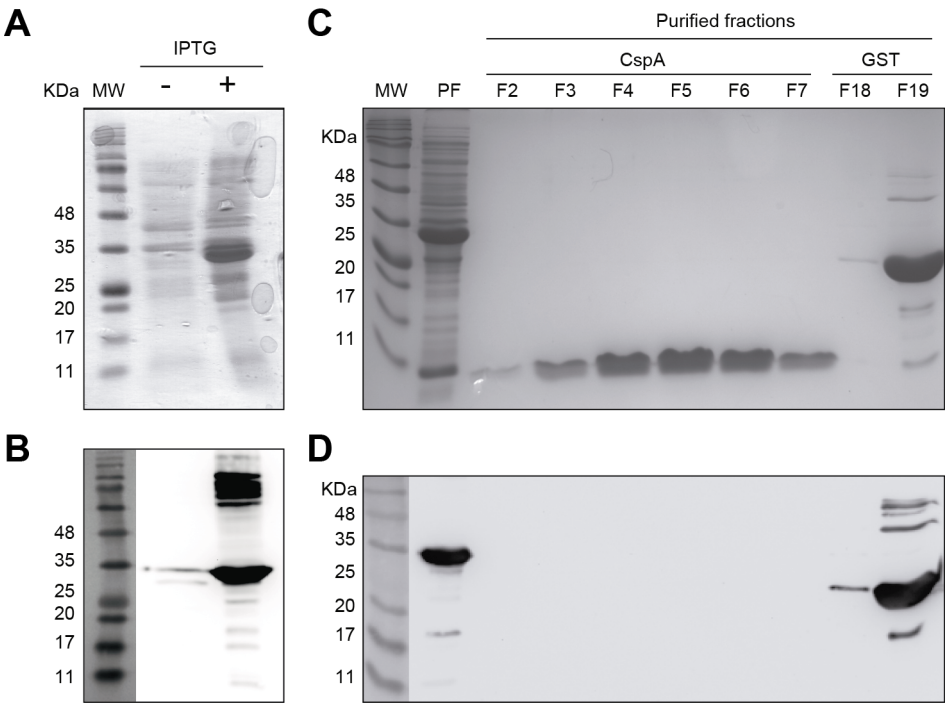

Figure 16. Purification of the recombinant CspA protein. A. Coomassie stain of a representative acrylamide gel showing the expression of GST-CspA fusion protein after before and after IPTG induction. **B.** Western blot performed with the samples used in A. GST-CspA fusion protein was developed using anti-GST antibodies. **C.** Coomassie stain of a representative acrylamide gel showing some of the purified fractions after digestion of GST-CspA fusion protein with Prescission protease. **D.** Western blot performed with the samples used in B, showing that the fractions containing the processed CspA lacked contamination with GST fragment. GST protein was developed using anti-GST antibodies.

Having purified rCspA, we performed EMSAs with the whole *cspC* mRNA, which we synthesized *in vitro*, radioactively labelled the 5'-end with and incubated with increasing concentrations of rCspA. As shown in Figure 17A, rCspA dependent mRNA band shifts are noticeable, indicating that rCspA and the *cspC* mRNA effectively interacted.

To determine what region of the *cspC* mRNA was recognized by CspA, we designed ssDNA oligonucleotides that covered the entire mRNA molecule. The reason for using ssDNA was because CSPs bind ssDNA as efficiently as RNA molecules and are more stable than RNA oligonucleotides (Phadtare and Inouye, 1999). Each ssDNA oligonucleotide overlapped approximately 10-nt with the following one (Fig. 17B). In this manner, we avoided missing any possible interactions at the extremes of the oligonucleotides. We radioactively labelled ssDNA oligos at the 5'-end and incubated with increasing concentrations of rCspA. Then, we run the samples in non-denaturing 10% polyacrylamide gels and the results showed that CspA bound f-C1 oligonucleotide with the highest affinity when compared to than the other six (Fig. 17C). This indicated that CspA interacted with the 5'UTR of the *cspC* mRNA *in vitro*.

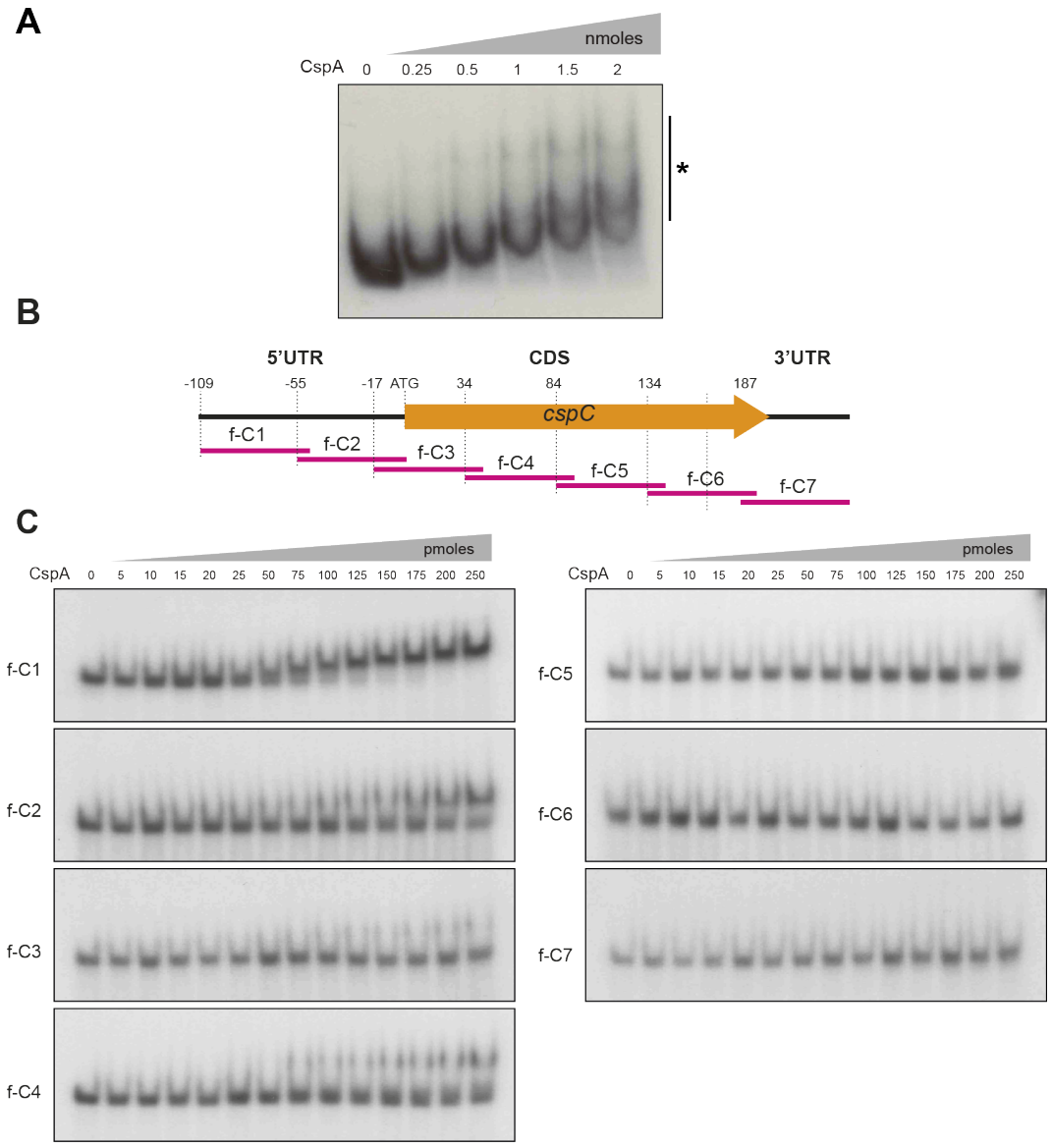

Figure 17. CspA binds the *cspC* mRNA at the 5'UTR *in vitro*. A. EMSA of CspA-*cspC* mRNA. Band-shifts originated from the combination of 20 fmol of ^{32}P-labeled synthetic *cspC* mRNA and increasing amounts of recombinant CspA (rCspA) (0.25 to 2 nmoles). **B.** Schematic representation of the ssDNA oligonucleotides designed to perform EMSAs with the rCspA protein. **C.** EMSA of f-C1 to f-C7 ssDNA oligonucleotides (20 fmol of ^{32}P-labelled synthetic oligo fragments) with increasing amounts of rCspA protein. The pmoles per reaction used in each lane are indicated. The CspA-oligonucleotide complexes are marked with an asterisk.

CspA represses CspC protein expression without affecting *cspC* mRNA levels *in vivo*

To further analyse the suppressor effect of CspA on CspC expression at the post-transcriptional level, we carried out Northern blots using RNA samples extracted from the WT and Δ*cspA* strains carrying the pCN51-cspC^{3xFLAG} plasmid and grown until mid-exponential phase. We developed the transferred RNAs using an antisense oligonucleotide that targeted the 3xFLAG sequence (as-3xFLAG) as a probe. This probe detected only the plasmidic *cspC*^{3xFLAG} mRNA, which was expressed under the control of the heterologous P*cad-cadC* module (Charpentier *et al.*, 2004). As shown in Figure 18, *cspC*^{3xFLAG} mRNA levels expressed from pCN51-cspC^{3xFLAG} were similar in the WT and Δ*cspA* strains, suggesting that CspA binding decreased CspC expression post-transcriptionally by impairing translation and without affecting *cspC* mRNA concentration.

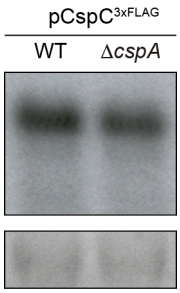

Figure 18. CspA binding does not affect *cspC* mRNA levels *in vivo*. Total RNA extraction was performed after growth until mid-exponential phase at 37ºC and 200 rpm. RNA samples were run into 1.25% agarose gels and transferred to Nitran membranes. Northern blots were developed using a ^{32}P-labelled anti-FLAG oligo probe and autoradiography. Ethidium bromide staining of 16S RNA is shown as loading control (LC). Protein samples were run into 12% polyacrylamide gels and transferred to Nitrocellulose membranes.

CspA inhibits CspC translation *in vitro*

Previous results suggested that CspA bound the 5'UTR region of *cspC* mRNA. Since such binding did not affect mRNA levels, differences on CspC levels between the Δ*cspA* and WT strain, could be explained by a direct alteration of CspC translation. To confirm that CspC regulation by CspA binding was uncoupled from transcription and mRNA turnover, we performed *in vitro* translation of a synthetic mRNA encoding the 3xFLAG-tagged CspC protein, in the presence and absence of rCspA. We used BSA protein as a negative control. Western blot results revealed that $CspC^{3xFLAG}$ could not be efficiently produced *in vitro* in the presence of rCspA (Fig. 19). This result was in agreement with the idea that CspA may be acting as a repressor of translation.

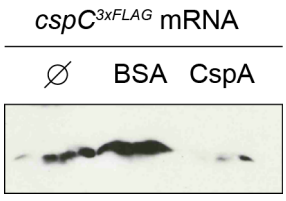

Figure 19. CspA inhibits CspC translation *in vitro*. Western blot of $CspC^{3xFLAG}$ protein after *in vitro* translation of the synthetic $cspC^{3xFLAG}$ mRNA in presence of BSA (lane 2) or recombinant CspA proteins (lane 3). Lane 1 shows the result from the *in vitro* translation of the $cspC^{3xFLAG}$ mRNA alone as a positive control.

CHAPTER IV:
Insights into the autoregulation of CspA expression

CHAPTER IV:

Insights into the auto-regulation of CspA expression.

In the previous chapter, we showed that CspA could bind the 5' UTR of *cspC* mRNA and act as direct repressor of CspC translation without affecting its mRNA levels. We wondered whether this was a generalized mode of action extended to all its negatively-regulated targets. For this reason, we decided to deeply analyse the auto-regulation of CspA expression (revealed in Chapter II) as an additional example of CspA-mediated repression.

CspA affects its own mRNA processing and protein expression

To understand how CspA affects its own expression at the post-transcriptional level, we performed Northern blots and compared the expression of the plasmidic *cspA*^{3xFLAG} mRNA in the WT and Δ*cspA* strains using a radioactively labelled antisense 3xFLAG oligonucleotide. This probe allowed us to specifically detect the *cspA*^{3xFLAG} mRNA expressed from the pCN51-cspA^{3xFLAG} plasmid, which is under the control of the heterologous P*cad-cadC* module (Charpentier *et al.*, 2004). In contrast to what was observed for CspC, Northern blots revealed two

cspA^{3xFLAG} mRNA bands in the WT and Δ*cspA* strains, indicating the presence of a processing site (Fig. 20). Such processed mRNA may be due to RNase III activity, as previously described (Lioliou *et al.*, 2012) (Fig. 20A). In agreement with this idea, the lower band in samples from the Δ*rnc* and Δ*rnc*Δ*cspA* strains, carrying the pCspA^{3xFLAG} plasmid, was absent (Fig. 20B). Interestingly, although the levels of the non-processed mRNA were only slightly higher in Δ*cspA* compared to the the WT strain, the amount of the processed mRNA in the former was significantly elevated (Fig. 20C). Lioliou and colleagues suggested that such processed *cspA* mRNA would be more prone to be translated (Lioliou *et al.*, 2012). Our observations indicated that CspA might interfere with RNase III processing and impair its own translation, which correlated with the detected higher CspA^{3xFLAG} protein levels in the Δ*cspA* strain (Fig. 20D). Western blots also revealed that CspA^{3xFLAG} levels, expressed from the pCspA^{3xFLAG} plasmid, were lower in the Δ*rnc*Δ*cspA* double mutant than in the Δ*cspA* single mutant strain, supporting that RNase III processing is participating in CspA auto-regulation (Fig. 20D). However, CspA^{3xFLAG} levels in the Δ*rnc*Δ*cspA* double mutant were slightly higher than in the WT and Δ*rnc* strains, indicating the existence of an additional cooperative RNase III-independent mechanism contributing to CspA repression (Fig. 20D). Altogether, these results suggested that CspA modulated its expression by, at least, interfering with RNase III processing of its own mRNA.

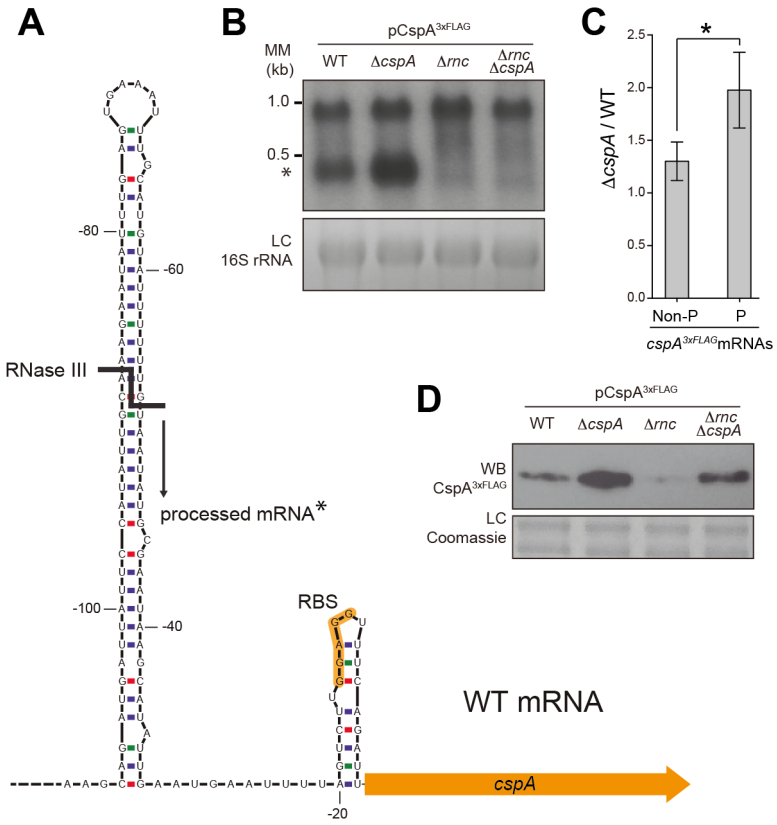

Figure 20. Processing of the *cspA* mRNA by RNase III. A. Schematic representation of the *cspA* 5'UTR stem loop adpated from Lioliou and colleagues (Lioliou *et al.*, 2012). The RNase III processing site is represented by a bold line and the processed mRNA is indicated with an asterisk. **B.** *cspA*^{3xFLAG} mRNA levels produced from the pCspA^{3xFLAG} plasmid in the WT, Δ*cspA*, Δ*rnc* and Δ*rnc*Δ*cspA* strains. Total RNA extraction was performed after growth until mid-exponential phase at 37ºC and 200 rpm. RNA samples were run into 1.25% agarose gels and transferred to Nitran membranes. Northern blots were developed using a ^{32}P-labelled anti-FLAG oligo probe and autoradiography. Ethidium bromide staining of 16S RNAs is shown as loading control (LC). MM, Millenium Marker. **C.** Comparison of non-processed and processed *cspA*^{3xFLAG} mRNA levels in the Δ*cspA* and the WT strain. Bands were quantified by densitometry of Northern blot auto-radiographies using Fiji (https://fiji.sc/). The mean ratios and standard deviations were calculated from four independent experiments. The asterisk indicates the statistical significance of the ratio differences (P = 0.0154). **D.** CspA^{3xFLAG} protein levels from the same strains and conditions described in B. Protein samples were run into 12% polyacrylamide gels and transferred to Nitrocellulose membranes. Western blots were developed using peroxidase conjugated anti-FLAG antibodies and a bioluminescence kit. Coomassie stained gel portions are shown as loading controls (LC).

The *cspA* 5'UTR is required for CspA expression and auto-regulation

The hairpin structure targeted by RNase III is located in the 5'UTR of the *cspA* mRNA (Lioliou *et al.*, 2012). To verify whether CspA was acting on the 5'UTR, we generated the pCspA^{3xFLAG}Δ5'UTR plasmid. This construct carried a version of the *cspA*^{3xFLAG} mRNA that lacked most of the 5'UTR while preserving the RBS (up to -18 nt from the start codon). We transformed the WT and Δ*cspA* strains with pCspA^{3xFLAG}Δ5'UTR and then addressed their *cspA*^{3xFLAG} mRNA expression by Northern blot. As anticipated, only one mRNA band with a similar intensity was found in both strains, indicating that the presence of the 5'UTR RNA was needed for RNase III processing (Fig. 21A) (Lioliou *et al.*, 2012). Western blot experiments showed that deletion of the 5'UTR drastically decreased CspA production in the WT and Δ*cspA* strains, nonetheless, the amount of protein was comparable between them (Fig. 21B). Altogether, these results indicated that mRNA processing is essential for an appropriate CspA expression and that the CspA auto-regulatory mechanism requires the 5'UTR of the *cspA* mRNA.

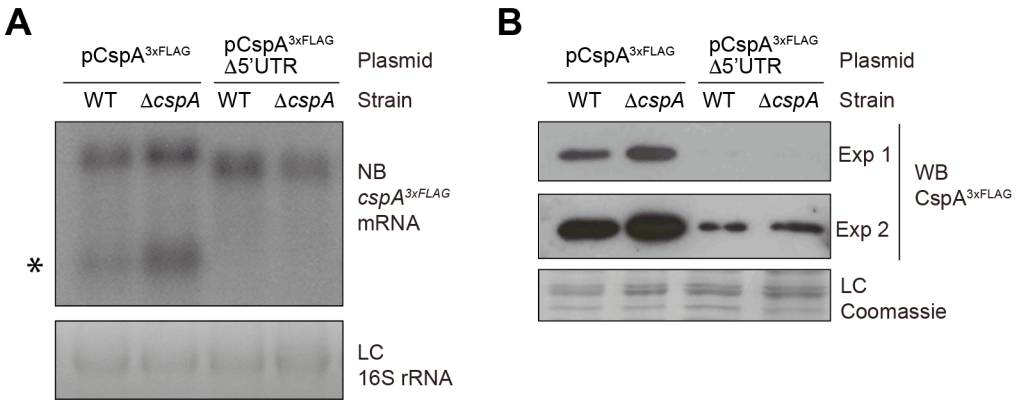

Figure 21. CspA expression and auto-regulation require the *cspA* 5'UTR. A. *cspA*^{3xFLAG} mRNAs generated from pCspA^{3xFLAG} and pCspA^{3xFLAG}-Δ5'UTR plasmids in the WT and Δ*cspA* strains. RNAs were extracted, run, transferred and developed as described in Figure 20. Ethidium bromide staining of 16S RNA is shown as a loading control (LC). Processed *cspA*^{3xFLAG} mRNA is indicated with an asterisk. **B.** CspA^{3xFLAG} protein levels from the pCspA^{3xFLAG} and pCspA^{3xFLAG}-Δ5'UTR plasmids in the WT and Δ*cspA* strains. Proteins were extracted, run, transferred and developed as described in Figure 20. Exp 1 and Exp 2 indicate two different exposure times. A Coomassie stained gel portion is shown as a loading control (LC).

CspA might bind a U-rich region located in the right arm of the 5'UTR stem-loop

In order to find which region of the *cspA* 5'UTR was bound by CspA, we designed three single-stranded DNA (ssDNA) oligonucleotides of about 60-nt long that completely covered the 5'UTR and part of the *cspA* CDS. As described in Chapter III, we used ssDNA oligos that overlapped 10 nt with each other (Fig. 22A) (Phadtare and Inouye, 1999). Following such strategy, f-A1 comprised the stretch of nucleotides from the TSS (-112) to position -51, f-A2 from -62 to +3 and f-A3 from -17 to +43 (considering position +1 the A of the start codon). After radiolabelling and incubating all three ssDNA oligonucleotides with increasing concentrations of purified recombinant CspA (rCspA), we run them in non-denaturing 10% polyacrylamide gels. Results showed that CspA bound f-A1 and f-A2 with dissociation constants (Kd) of approximately 1.5 and 2.8 μM, respectively. In contrast, CspA poorly bound f-A3 (Kd > 8.5 μM) (Fig. 22B). To demonstrate that CspA binding was specific, we performed competition assays using increasing concentrations of cold unlabelled f-A1 and f-A2. As expected, these oligonucleotides competed with the labelled ones (Fig. 22C). Additionally, the f-CDS oligonucleotide (from 148 to 201 of the CspA CDS sequence) was included as negative control (Fig. 22D). Altogether, these assays showed that CspA bound specifically and more efficiently

regions covered by f-A1 and f-A2 than those included by f-A3 and f-CDS. This implied that CspA either bound more than one site of the 5'-UTR or the 10-nt overlapping region, between f-A1 and f-A2. Coincidentally, such overlapping sequence included a thymidine-rich (T-rich) stretch (Fig. 22A). Uracil-rich (U-rich) regions of RNAs (or T-rich regions for ssDNA) have been previously proposed as potential targets for some CSPs (Phadtare and Inouye, 1999; Lopez *et al.*, 2001; Zeeb *et al.*, 2006; Max *et al.*, 2007; Sachs *et al.*, 2012; Lee *et al.*, 2013; Benhalevy *et al.*, 2015). Therefore, we wondered if the T stretch, common to both oligonucleotides, was involved in the band shifts observed in the presence of CspA. For this purpose, we repeated the EMSAs using modified f-A1 and f-A2 (f-A1-ΔT and f-A2-ΔT) that lacked the T-rich sequences from the 3'- and 5'-end, respectively. Results showed a remarkable decrease in the affinity of CspA for f-A1-ΔT (Kd > 8.5 μM) and f-A2-ΔT (Kd ~ 7 μM) compared to that for the original ssDNA oligonucleotides (Fig. 22E). This indicated that the T-rich region enhanced CspA binding and suggested a putative interaction between CspA and the U-rich region of *cspA* 5'UTR.

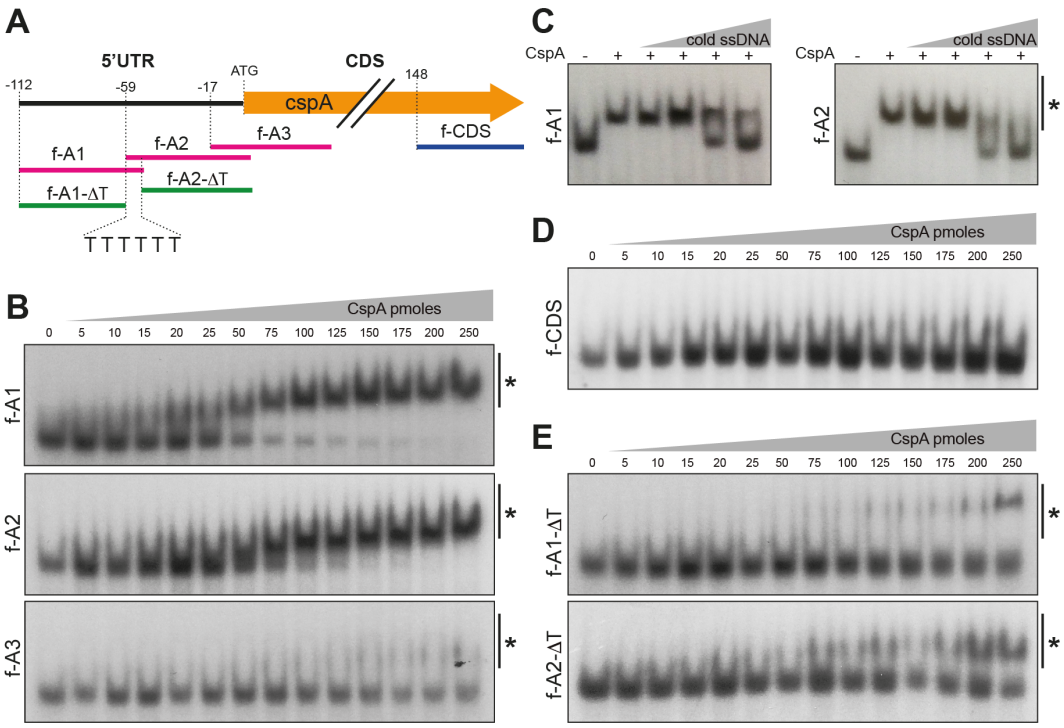

Figure 22. CspA binds to a T-rich motif *in vitro*. A. Schematic representation of the ssDNA oligonucleotides designed to perform the EMSAs with recombinant CspA (rCspA). **B.** EMSA of f-A1, f-A2 and f-A3 ssDNA oligonucleotides (20 fmol of ^{32}P-labelled synthetic oligo fragments) with increasing amounts of rCspA. The pmoles per reaction used in each lane are indicated. **C.** Gel shift competition assay of labelled f-A1 and f-A2 performed in the presence of increasing concentrations of unlabelled f-A1 and f-A2 and 25 and 50 pmoles of rCspA, respectively. **D.** EMSA of the f-CDS ssDNA oligonucleotide (20 fmol of ^{32}P-labelled synthetic oligo fragments) with increasing amounts of rCspA**. E.** EMSA of the f-A1-ΔT and f-A2-ΔT ssDNA oligonucleotides. These fragments lack the T-rich region from the 3' and 5' ends of f-A1 and f-A2, respectively. The rCspA-oligonucleotide complexes are indicated with an asterisk.

CspA melts the stem-loop of the *cspA* 5'UTR *in vitro*

The U stretch is located in the right arm of the stem-loop structure processed by RNase III (Fig. 20A). Based on this observation, we hypothesized that CspA may bind to this region, melt the RNA stem-loop and prevent the RNase III processing. To test this hypothesis, we used a molecular beacon system, as previously described (Phadtare, Inouye, *et al.*, 2002). In our case, it consisted of a 49-mer ssDNA oligonucleotide that included the 5'UTR stem-loop (from position -95 to -45 considering the A of the start codon as position +1), incorporating a molecule of fluorescein (FAM) and the Black Hole Quencher 1 (BHQ_1), which were attached to the 5'- and 3'-end of the oligonucleotide, respectively (Fig. 23A). If these two molecules fell in close proximity to each other, an efficient quenching of the FAM fluorescence would occur. To test the viability of the designed molecular beacon system, we measured the fluorescence levels at different temperatures. Figure 23B shows basal fluorescence levels at 37 and 45ºC, indicating that the molecular beacon was properly folded and retained a FAM quenched state due to the proximity of BHQ_1. In contrast, an appreciable increase of fluorescence levels occurred when this structure was incubated at temperatures higher than 55ºC, evidencing a melting process in which the FAM was separated from the quencher.

Since the molecular beacon was working as expected, we then incubated it with purified BSA, as a negative control, and rCspA for 10 min at 37ºC and measured the fluorescence. Addition of rCspA resulted in an increase of the beacon fluorescence compared to the BSA treatment, indicating that the interaction of rCspA with the stem loop structure eventually led to its disruption (Fig. 23C). To test if this effect was due to the presence of rCspA, Proteinase K was added to the mixtures and these incubated for 30 min at 37ºC. After the incubation period, fluorescence decreased as Proteinase K degraded rCspA (Fig. 23D). Finally, the mixtures were incubated for 10 min at 65ºC to melt the beacon structure. Both mixtures, containing rCspA and BSA, presented an increase in the fluorescence levels upon temperature shift, confirming that the beacon was still functional after the treatments (Fig. 23E). From these results, we learned that CspA was able to disrupt the ssDNA stem-loop structure of the *cspA* 5'UTR *in vitro*. The consequence of such disruption in the RNA molecule *in vivo*, would prevent RNase III from cleaving the *cspA* mRNA.

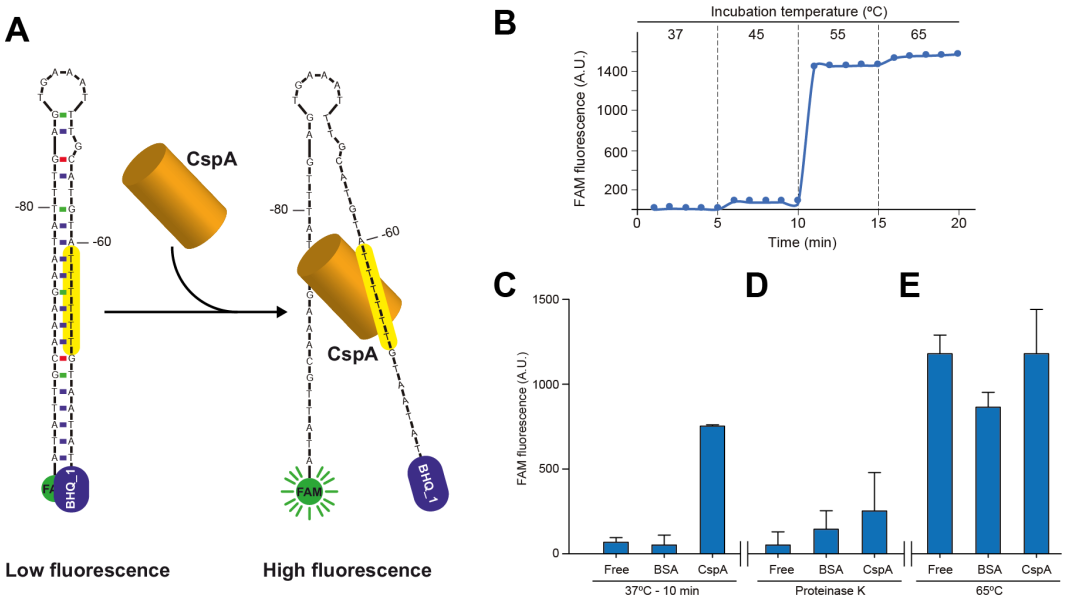

Figure 23. CspA melts the stem loop of the *cspA* 5'UTR *in vitro*. A. Schematic representation of the molecular beacon design that mimics the stem-loop of the *cspA* 5'UTR. The 5'-end and the 3'-end harbour fluorescein (FAM) and the Black Hole Quencher 1 (BHQ_1), respectively. The structure would keep FAM close to BHQ_1 avoiding fluorescence emission. The T-rich motif is indicated in yellow. **B.** Control test to monitor the effectiveness of the molecular beacon design. A temperature increase over 50ºC for 10 min caused melting of the molecular beacon secondary structure and led to fluorescence emission by FAM. **C.** FAM fluorescence after incubation of the molecular beacon with 7 nmoles of recombinant CspA (rCspA) or BSA for 10 min at 37ºC. BSA was introduced as a negative control. The presence of rCspA produced a similar effect to temperature increase by promoting fluorescence emission. **D.** FAM fluorescence levels after addition of 200 μg of Proteinase K to the previous mixtures and incubation for 30 min. Proteinase K processed rCspA and BSA, leading to re-naturalization of the beacon structure and, as consequence, a reduction in fluorescence emission. **E.** FAM fluorescence after treating the previous mixtures for 10 min at 65ºC. Raise in temperature caused an increase of fluorescence due to melting of secondary structures.

CspA requires the stem-loop containing the U-rich motif to regulate its own expression *in vivo*

Previous results suggested that CspA might interfere with the *cspA* 5'UTR processing by RNase III. At least, two key elements are necessary for this processing to occur: i) a correctly folded stem loop at the 5'UTR, generating a double-stranded RNA region, and ii) RNase III being able to effectively cleave it (Fig. 20A) (Lioliou *et al.*, 2012). There are several hypotheses that might explain the negative effect exerted by CspA on the RNase-III processing. CspA could either unfold the stem loop into a single stranded RNA or allosterically interfere with RNase III binding. These two possibilities would result in inhibition of RNase III activity by direct interaction with the *cspA* mRNA. Another consideration is that CspA could indirectly affect RNase III expression. Proteomic analysis revealed that RNase III protein levels were reduced in the Δ*cspA* strain (-2.1, P-value = 0.0004) (Annex I). However, this last observation seems in disagreement with the increase in RNase III-processing observed for the *cspA* 5'UTR in Δ*cspA* (Fig. 20B). Hence, an interference with the processing of the stem loop due to direct binding of CspA, as suggested by the *in vitro* assays, seemed the most reliable hypothesis.

To evaluate *in vivo* the role of the stem loop in CspA auto-regulation, we constructed the pCspA^{3xFLAG}-M5U plasmid, which carried an exchange of nucleotides -59[TTTTT]-54 by nucleotides -59[GACAG]-54, and transformed the WT and Δ*cspA* strains with it. Such mutation was designed to generate a *cspA*^{3xFLAG} mRNA that, on the one hand, would prevent the formation of the hairpin targeted by RNase III and, on the other hand, would avoid CspA binding to *cspA* 5'UTR. As anticipated, Northern blots revealed that the mutation of the U-rich motif avoided the regular RNase III processing occurring at the *cspA* 5'UTR (Fig. 24B). In this case, the levels of the CspA^{3xFLAG} protein were not remarkably affected when Δ*cspA* was compared to the WT strain. This result indicated that CspA auto-regulation may require the U-rich motif for CspA binding and/or the formation of the stem-loop to be digested by RNase III.

Finally, since the U-rich mutation affected two variables simultaneously (CspA binding and RNase III processing of the *cspA* mRNA), we generated a second plasmid, the pCspA^{3xFLAG}-M5UC, which carried an additional mutation that substituted nucleotides -87[AAGAA]-83 for -87[CTGTC]-83. The purpose was to compensate the previous one (pCspA^{3xFLAG}-M5U) and restore the RNA hairpin at the 5'UTR. The stem loop, lacking the U-rich motif, would still be recognised and processed by RNase III. This plasmid was used to transform the WT and Δ*cspA* strains. Northern blots confirmed that the RNase III dependent mRNA processing was restored when *cspA* was transcribed from pCspA^{3xFLAG}-M5UC (Fig.

24C), although to a lower extent than that obtained from the mRNA generated from pCspA^{3xFLAG}. This observation could be explained by the location of the mutation, at the cleavage site of RNase III, which may decrease the processing efficiency of the mRNA (Pertzev, 2006). Nonetheless, the levels of CspA^{3xFLAG} protein were not notably different between the Δ*cspA* and WT strains, indicating that the U-rich region is required for CspA repression (Fig. 24). Overall, these observations illustrated the existence of a very complex regulatory network where the 5'UTR stem-loop is crucial in the control of CspA expression. Interestingly, this structure is alternatively targeted by two RBPs, RNase III and CspA itself.

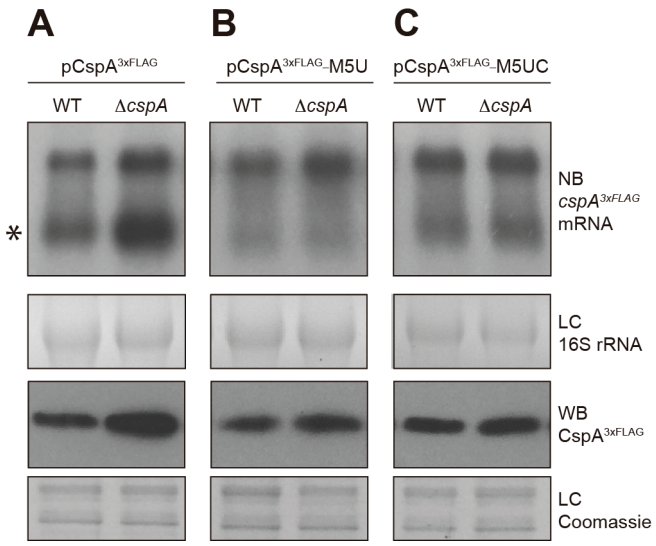

Figure 24. The U-rich motif is required for CspA auto-regulation *in vivo*. mRNA and protein levels of CspA^{3xFLAG} expressed from pCspA^{3xFLAG} (**A**), pCspA^{3xFLAG}-M5U (**B**) and pCspA^{3xFLAG}-M5UC (**C**) in the WT and Δ*cspA* strains. RNAs and proteins were extracted, run, transferred and developed as indicated in Fig. 20. In the Northern blots, the processed *cspA*^{3xFLAG} mRNA, indicated with an asterisk. Ethidium bromide staining of 16S RNA is shown as loading control (LC). Western blots include Coomassie stained gel portions as loading controls (LC).

DISCUSSION

DISCUSSION

Several RBPs with unrelated protein domains have been proposed to act as global post-transcriptional modulators of gene expression (Van Assche *et al.*, 2015). Up to now, the targetomes of some regulatory RBPs have been unveiled (e.g. RNA chaperones Hfq, CsrA, ProQ, CSPs and ribonucleases RNase III, RNase E). In all cases, hundreds of RNA targets comprising both coding and non-coding RNAs were found, showing that RBPs constitute central elements in fine-tuning bacterial biological processes (Zhang *et al.*, 2003; Sittka *et al.*, 2008; Lioliou *et al.*, 2012; Chao *et al.*, 2012; Dambach *et al.*, 2013; Holmqvist *et al.*, 2016; Smirnov *et al.*, 2016; Chao *et al.*, 2017; Michaux *et al.*, 2017). Most of these studies have been conducted in Gram-negative bacterial models, leaving Gram-positives with only a few examples. Throughout this Thesis, different functional aspects of the RNA chaperone CspA have been analysed using *S. aureus,* a Gram-positive pathogen of major clinical relevance, as a model. Here, we propose a different functionality between *S. aureus* CspA and its paralogs, CspB and CspC. In addition, we suggest that CSPs functional specificity may be encoded in the carboxi-half of the protein. We also provide the regulon of CspA and insights on the mechanisms by which CspA negatively regulates the expression of CspC and itself.

S. aureus CSPs are not functionally redundant

Based on their sequence identity, it has been long speculated that CSP paralogs present in a genome may share the same biological function. A recent report by Michaux and colleagues, demonstrated that CspC and CspE RNA targets from *Salmonella* overlap in a large percentage (Michaux *et al.*, 2017). Simultaneous mutation of *cspC* and *cspE* genes resulted in strong phenotypic changes. However, single-gene deletions did not impair the analysed phenotypes indicating redundant functions for these two proteins (Michaux *et al.*, 2017). Likewise, both CspC and CspE in *E. coli* are required to resist environmental stresses by stabilizing the *rpoS* transcript at the entry of stationary phase (Shenhar *et al.*, 2012). Although these findings show that CSPs may have redundant functions, it is important to highlight that *Salmonella* and *E. coli* encode six and nine CSPs, respectively. Besides, all the remaining CSPs appear to have different biological roles (Xia *et al.*, 2001; Michaux *et al.*, 2017). This suggests that a few amino acids can dictate CSP target specificity and ultimately define particular biological functions for them. Our research supports this notion showing that, although *S. aureus* CSPs display a very close sequence identity, CspA function could not be replaced by paralogs CspB and CspC *in vivo* (Fig. 6). Generation of chimeric CSPs indicated that the elements responsible for target specificity may be encoded in the

carboxi-half region of the protein, comprising in β-strands 4 and 5 and excluding the amino acidic loop connecting motifs RNP1 and RNP2 (Fig. 8). When comparing *S. aureus* CSP protein sequences, and more specifically the carboxi-half we found position 58 to be different for all three paralogs. In the case of CspA it was a proline (P58), which seemed relevant since it was located in β-strand 4 and, from a three-dimensional perspective, in front of RNP1 and RNP2. Moreover, P58 is conserved in CspC and CspE of *E. coli* and in CspB of *B. subtilis* (BsCspB), which shares a 76.9% of identity with *S. aureus* CspA. For this reason, we constructed a plasmid expressing BsCspB to complement the *S. aureus cspA* mutant strain. Unfortunately, STX production could not be restored (data not shown), indicating that probably a combination of amino acids is what determines CSP functional specificity. We are currently performing new chimeras that might help us identifying such amino acids.

In summary, our results are in agreement with the idea of independent regulons for each of *S. aureus* CSPs. However, identifying the amino acids that dictate CspA differential binding as well as CspB and CspC RNA targets is a future requirement for understanding CSPs target specificity.

CspA is a global regulator in *S. aureus*

The regulon of *S. aureus* CspA unveiled here includes more than two hundred potential direct RNA targets, presenting it as a global post-transcriptional regulator. The data was loaded into a public web server (http://rnamaps.unavarra.es/), which allows any user to quickly look for CspA-binding signals on selected genes.

Restricting the identification of CspA targets to only those appearing in both peak-calling methods (thresholding by TAS and CisGenome) may somehow have underestimated the number of CspA targets shown in Annex II. Besides, the proteomic analysis was only able to detect about 45% of *S. aureus* proteins, making it reasonable to argue that CspA possibly controls the expression of more proteins than those presented here (Annex I). Nevertheless, it is noteworthy that several of the targets that we identified, such as the alternative sigma factor B (SigB), the DEAD-box RNA helicase, specific ribosomal proteins, superoxide dismutase and other stress-associated proteins, have been previously related to CSP regulation in other bacteria (Tanaka *et al.*, 2012; Michaux *et al.*, 2012).

We showed that deletion of the *cspA* gene produced significant changes on the protein expression of many genes (Annex I) that could be correlated with relevant phenotypic changes. Increased bacterial

aggregation, most likely due to an increase in PNAG biosynthesis, and impaired resistance to oxidative stress are good examples of this (Fig. 11). These observations indicated that CspA is required for *S. aureus* to adapt to different environmental niches. It also supports the idea that CspA may recognise targets from specific functional groups, as the GO term enrichment analysis revealed (Fig. 10). This made us wonder if such bias was also true sequence-wise, by recognizing specific RNA nucleotide patterns. However, analysis of CspA binding regions showed that CspA had no preference for a particular region among the putative targets. Failing to find a consensus sequence for CspA may be explained by the fact that cold shock domains bind short nucleotide regions, which have been estimated to be around 5-7 nt (Lopez *et al.*, 1999; Sachs *et al.*, 2012). It could also be speculated that the position of the peaks detected by peak-calling may not correspond to the initial binding sites of CspA but just be indicative of the region where CspA was placed at the moment of the RIP assay. In agreement with this, the position of the binding peak on the *cspC* mRNA did not correlate with the results of the EMSAs (Fig 15A and 17C). Additionally, since sRNAs were identified in the RIP-chip analysis, we cannot exclude that some of the peaks were indirectly originated by pulling down mRNA targets of those sRNAs. Regardless, categorization of the differentially expressed targets suggested that CspA binds to transcripts that may be functionally related (Fig. 10). Thus, there

might be a common pattern (RNA sequence, RNA structure or a combination of both) for CspA targets that still needs to be identified.

Based on previous studies, it is reasonable to expect regulatory elements that specifically impair translation efficiency in CspA up-regulated targets (Kudla *et al.*, 2009; Pop *et al.*, 2014). Recently, deep sequencing-based ribosome profiling analysis have been performed in *S. aureus* (Davis *et al.*, 2014; Basu and Yap, 2016). Ribosome-protected footprints (RPFs) placed at certain positions can be an indicator of ribosome-stalling sites. In contrast, RPFs distributed across a particular transcript implies that the mRNA is actively translated. To check if CspA targets contain stalling signals, we examined the already available RPF data (Davis *et al.*, 2014; Basu and Yap, 2016). Interestingly, we found that several positively regulated CspA targets contained RPFs signals, which could represent putative ribosome-stalling sites. To gain a deeper insight, we looked for the presence of stem-loop structures at the RPF locations using the Mfold software (Zuker, 2003). Some of the targets contained predicted secondary structures in RPF surroundings (Fig. 25). Given such coincidence, one could hypothesize that *S. aureus* CspA may enhance translation by disrupting the mentioned secondary structures, an idea that was previously explored in other bacteria like *E. coli* in cold shock conditions (Phadtare and Severinov, 2010).

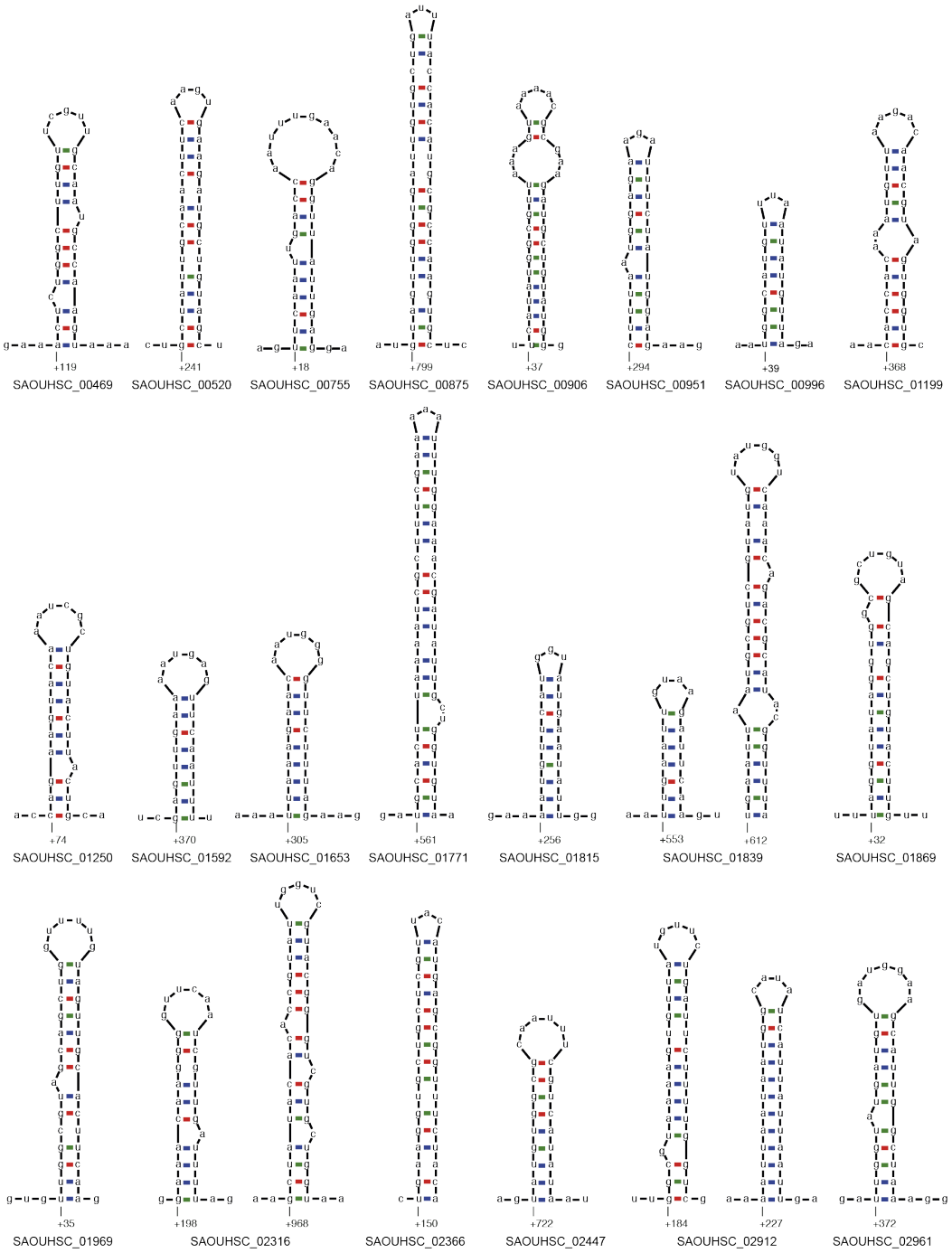

Figure 25. Putative ribosome stalling structures. M-fold predictions of RNA secondary structures in CspA directly regulated mRNAs that presented RFPs, as previously described (see text for details).

Since *S. aureus* CspA cannot be considered as part of the cold shock regulon (it is a non-cold-induced CSP), it would rather act on mRNA structures that are formed at the optimal growing temperature. Nevertheless, further investigations are required to demonstrate that the RNA secondary structures present in CspA targets are able to stall ribosomes, impair translation *in vivo* and susceptible to melting by interaction with CspA.

Considering this idea, ribosome-profiling maps in the Δ*cspA* strain would prove useful for verifying an enrichment in RPFs in those CspA targets whose expression is decreased in its absence. Alternatively, CspA binding might impair mRNA degradation by changing the susceptibility to ribonucleases and, as a result, increase translation (Bonnin and Bouloc, 2015).

CspA is able to repress gene expression by different mechanisms

The ability of CspA to reduce the expression of some of its putative targets opened the possibility to new regulatory mechanisms other than translation enhancement. For this reason, we focused on analysing more extensively how CspA binding could repress gene expression.

In the case of CspC, the post-transcriptional repression did not imply variations in the mRNA levels (Fig. 18), indicating that the molecular mode of action might be related to the translation process. This was supported

by *in vitro* translation experiments, which showed that the *cpsC*^{3xFLAG} mRNA was poorly translated in the presence of CspA. It could be speculated that CspA interferes with the translation process by binding the 5'UTR, as the EMSAs indicated (Fig. 17), and blocks the access to the ribosome. However, further experiments are required to demonstrate the precise mechanism of regulation.

Regarding CspA repression, the mode of action seems to have different implications than those observed for CspC. Here, we propose an auto-regulatory mechanism in which a U-rich motif, located in the hairpin of *S. aureus cspA* 5'UTR, seems of importance for CspA to bind its own mRNA and regulate its own expression. In principle, the interaction between CspA and the *cspA* mRNA would either disrupt the RNA loop or allosterically avoid RNase III binding. The molecular beacon experiment suggests that the first possibility is true (Fig. 23). However, since CspA targets the RNase III cleavage site, the second possibility cannot be totally excluded. Regardless of the precise mechanism of action, the consequence would be a lack of processing of the *cspA* mRNA, favouring a less translated form. In contrast, a lack of interaction between CspA and the stem loop, would allow RNase III to cleave it and generate a shorter *cspA* mRNA. The resulting processed mRNA would present a more efficiently translated conformation (Lioliou *et al.*, 2012). In summary, CspA would also be acting as a translation inhibitor for its own mRNA but through a different mechanism to the one adopted for regulating CspC.

Model of CspA auto-regulation

One could easily imagine the CspA auto-regulagtory mechanism as a way to sense the intracellular levels of CspA, in which sufficient concentration of the protein would eventually trigger the proposed mechanism (Fig. 26). This hypothetical model is in agreement with the results shown in Figure 24 that demonstrate how the mutations of the *cspA* 5'UTR, which eliminated the U-rich motif but restored the stem-loop, did not significantly affect CspA expression. Remarkably, the U stretch is located in the denominated proximal-box (PB) of the hairpin, affecting the catalytic activity of RNase III (Pertzev, 2006).

The fact that $CspA^{3xFLAG}$ protein levels in the Δ*rnc*Δ*cspA* double mutant were slightly higher than in the WT and Δ*rnc* strains indicated that CspA might also repress its own expression in the non-processed mRNA (Fig. 20D). Although further investigation is needed, this might act as an additional mechanism to ensure CspA repression. For this reason, the presence of other factors in the *cspA* locus, such as another promoter located upstream of the precedent CDS (which generates a bicistronic transcript that includes *cspA*) and an antisense RNA, must also be considered (Lioliou *et al.*, 2012; Sahukhal and Elasri, 2014; Uppalapati *et al.*, 2017). In this study, we only contemplated the *cspA* monocytronic

transcript as a template for designing plasmid constructs since, according to our transcriptomic data (http://rnamaps.unavarra.es/), it was the predominant form. While this was sufficient to show a relevant self-regulatory role for CspA through its 5'UTR, we cannot exclude that CspA may also interact with the other transcripts mentioned above. Overall, these evidences highlight that a refined regulation is of importance for maintaining the proper levels of one of the most abundant staphylococcal proteins.

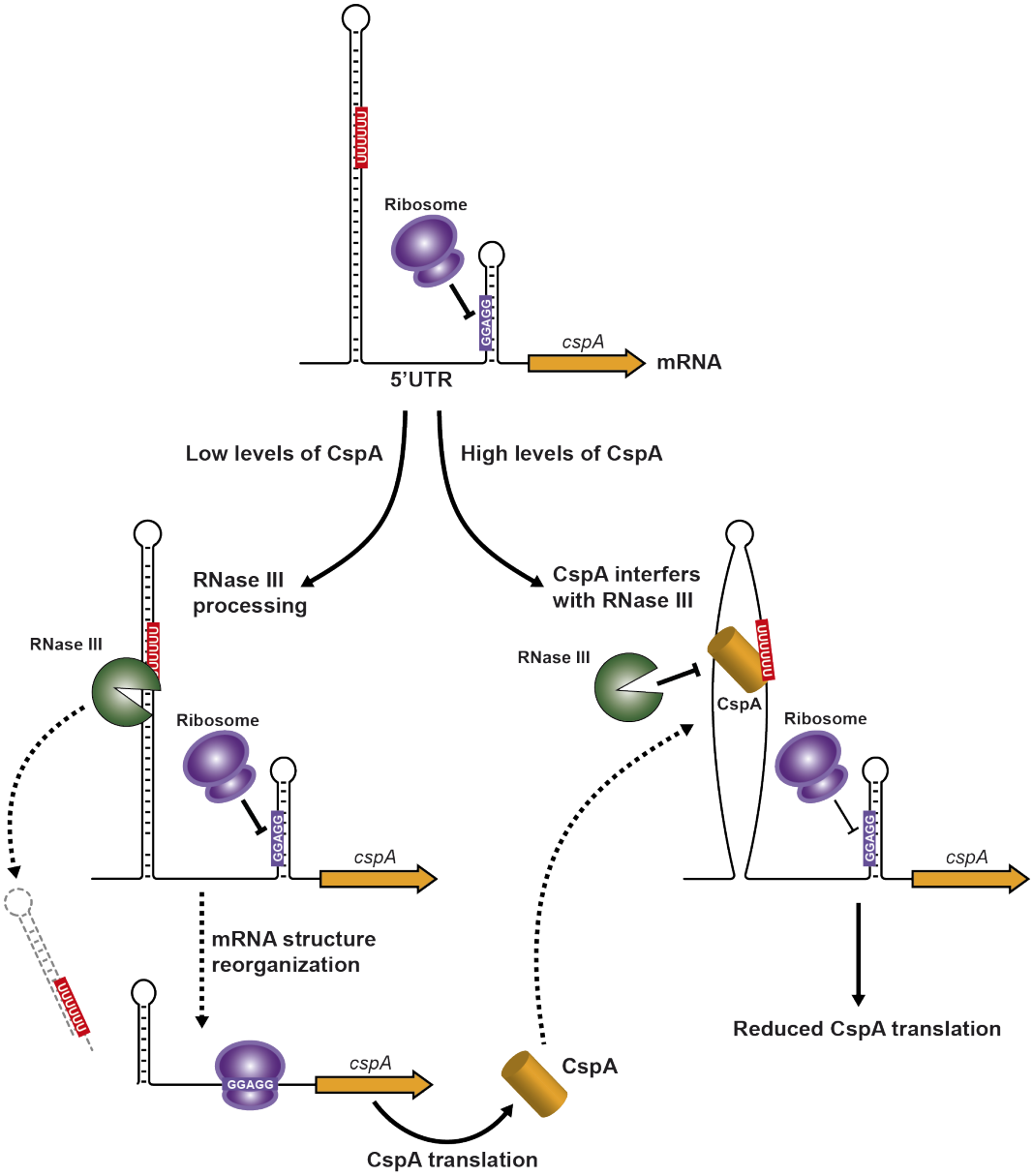

Figure 26. Model of the putative CspA auto-regulatory mechanism. Representation showing the stem-loop at the 5'UTR of *cspA* mRNA, which is targeted by RNase III. If newly transcribed *cspA* mRNA encounters low levels of CspA protein, RNase III can process its hairpin. The processed *cspA* mRNA may change its conformation allowing CspA translation. In contrast, if the levels of CspA protein are sufficient, it can target the U-rich region from the 5' UTR hairpin and disrupt the secondary structure. As a result, RNase III processing is compromised, hindering CspA translation. This model proposes a regulatory mechanism in which CspA acts as an antagonist element of the RNase III activity by melting the RNA double-stranded structure targeted by it.

Could CspA be an antagonist of RNase III activity?

The ability of CspA to interfere with RNase III activity made us wonder if the observed antagonist effect of both proteins could be a generalized mechanism. Lioliou and colleagues have already mapped the RNAs recognised by RNase III in *S. aureus* (Lioliou *et al.*, 2012). Therefore, we correlated the list of targets of both proteins and found that RNase III processed up to 80% of the mRNAs bound by CspA. Such a high overlap between both datasets suggested a putative antagonist function between CspA and RNase III, in which the former would act by disrupting the RNA structures targeted by the latter. Although this hypothesis seems plausible it deserves further investigation before making it a general assumption.

Regulatory specialization through specific RNA-elements and RBPs

In summary, the differences in both repressive regulatory processes supported the idea that post-transcriptional control of gene expression depends on the specific regulatory RNA elements located in each mRNA. Such elements can be modified by RBPs, like CspA, and consequently translation and/or stability of mRNA might change. The presence of different RNA regulatory elements could explain why bacteria need

several copies of different RBPs (e.g. many bacterial genomes contain more than one variant of CSPs, RNA helicases, S1 proteins).

From a general point of view, some paralogs may have the same biological function, however, the fact that most of them recognize different RNA targets shows a certain specialization for CSPs in bacteria. Discovering that RNA chaperones like *S. aureus* CspA target RNA structures recognized by other RBPs, opens new ways for understanding how CSPs modulate protein expression. This suggests that the nature of the RNA regulatory elements present in each gene and the way in which the different RBPs interact with them are key factors that determine the fate of protein expression and, ultimately, safeguard the correct development of every organism.

CONCLUSIONS

CONCLUSIONS

1) *S. aureus* CSPs share a high percentage of identity. However, despite being expressed in comparable amounts, only CspA is able to complement staphyloxanthin production in a Δ*cspA* strain, indicating different biological functions for each of them.
2) Complementation studies of the staphyloxanthin production in a Δ*cspA* strain, using different CSP chimeric proteins, suggests that the amino acidic differences in the carboxi-half of the proteins might determine the functional specificity of *S. aureus* CSPs.
3) *S. aureus* CSPs expression is differentially regulated at the post-transcriptional level, possibly through their untranslated regions.
4) Identification of the CspA regulon portrays this chaperone as a global gene expression regulator, mainly of genes related to metabolism, virulence and stress response. Therefore, deletion of the *cspA* gene generates evident phenotypic variations such as an increase in cellular aggregation and a defect in coping with oxidative stress agents.
5) CspA represses CspC translation upon binding the 5'UTR of the *cspC* mRNA, possibly by masking the RBS and without affecting its mRNA levels.

6) CspA represses its own expression by binding a U-rich motif, located in a stem-loop of its 5'UTR. This RNA structure is also targeted and cleaved by RNase III in order to improve mRNA translation.

7) The interaction of CspA with a molecular beacon, which mimics the targeted stem-loop, increases fluorescence emission. This indicates that CspA is be able to disrupt such double-stranded structure *in vitro* and, possibly, prevent RNase III processing *in vivo*.

8) Mutations that eliminate the U-rich motif while preserving the secondary structure of the *cspA* 5'UTR stem-loop, confirm that such motif is necessary for CspA auto-regulation *in vivo.*

9) The interaction of CspA with a stem-loop, which is also targeted by RNase III, suggests a functional antagonism between both RBPs that could apply to other RNase III targets.

CONCLUSIONES

CONCLUSIONES

1) Las CSPs de *S. aureus* tienen un alto porcentaje de identidad entre ellas. Sin embargo, y a pesar de expresarlas en cantidades similares, sólo CspA es capaz de complementar la producción de estafiloxantina en una cepa Δ*cspA,* indicando así una diferencia funcional entre las mismas.
2) Los estudios de complementación de la producción de estafiloxantina en una cepa Δ*cspA*, con distintas quimeras de CSPs, sugieren que las diferencias amino acídicas localizadas en la mitad carboxi-terminal de la proteína determinan la especificidad funcional de las CSPs de *S. aureus*.
3) La expresión de las CSPs de *S. aureus* se encuentra diferencialmente regulada a nivel post-transcripcional, probablemente a través de sus regiones no traducidas.
4) La identificación del regulón muestra a CspA como un regulador global de la expresión en *S. aureus*, principalmente de genes relacionados con el metabolismo, la virulencia y la respuesta al estrés. Por ello, la deleción del gen produce cambios fenotípicos evidentes como un aumento en la agregación celular y una disminución en su capacidad para hacer frente a agentes oxidantes.

5) CspA reprime la traducción de CspC tras unirse a la 5'UTR del mRNA de *cspC,* posiblemente ocultando la zona de unión al ribosoma, sin afectar a los niveles de dicho mRNA.
6) CspA reprime su propia expresión mediante la unión a un motivo rico en uridinas (U-rich), localizado en una horquilla de RNA de su propia 5'UTR. Esta horquilla es diana para el procesamiento del mRNA de *cspA* por RNase III, que es necesario para mejorar su traducción.
7) La interacción de CspA con una baliza molecular que imita a la horquilla de RNA, produce un aumento en la emisión de fluorescencia. Esto indica que CspA deshace esta estructura de doble cadena *in vitro* y posiblemente pueda evitar el corte por RNase III *in vivo*.
8) Las mutaciones que eliminan el motivo U-rich, pero que preservan la estructura de la horquilla presente en la 5'UTR de *cspA,* confirman que dicho motivo es necesario para la auto-regulación de la expresión de CspA *in vivo.*
9) La interacción de CspA con una estructura de RNA, que también es reconocida por RNase III, sugiere un antagonismo funcional entre ambas RBPs que podría ser extensible a otras dianas de RNasa III.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abeysirigunawardena, S.C., Kim, H., Lai, J., Ragunathan, K., Rappé, M.C., Luthey-Schulten, Z., *et al.* (2017) Evolution of protein-coupled RNA dynamics during hierarchical assembly of ribosomal complexes. *Nature Communications* **8**: 492.

Anderson, K.L., Roberts, C., Disz, T., Vonstein, V., Hwang, K., Overbeek, R., *et al.* (2006) Characterization of the *Staphylococcus aureus* Heat Shock, Cold Shock, Stringent, and SOS Responses and Their Effects on Log-Phase mRNA Turnover. *J Bacteriol* **188**: 6739–6756.

Apirion, D., and Lassar, A.B. (1978) A conditional lethal mutant of *Escherichia coli* which affects the processing of ribosomal RNA. *J Biol Chem* **253**: 1738–1742.

Arnaud, M., Chastanet, A., and Debarbouille, M. (2004) New Vector for Efficient Allelic Replacement in Naturally Nontransformable, Low-GC-Content, Gram-Positive Bacteria. *Appl Environ Microbiol* **70**: 6887–6891.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**: 195–201.

Aseev, L.V., Levandovskaya, A.A., Tchufistova, L.S., Scaptsova, N.V., and Boni, I.V. (2008) A new regulatory circuit in ribosomal protein operons: S2-mediated control of the rpsB-tsf expression in vivo. *RNA* **14**: 1882–1894.

Awano, N., Rajagopal, V., Arbing, M., Patel, S., Hunt, J., Inouye, M., and Phadtare, S. (2010) Escherichia coli RNase R has dual activities, helicase and RNase. *J Bacteriol* **192**: 1344–1352.

Babitzke, P., and Romeo, T. (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Current Opinion in Microbiology* **10**: 156–163.

Bae, W., Jones, P.G., and Inouye, M. (1997) CspA, the major cold shock protein of *Escherichia coli*, negatively regulates its own gene expression. *J Bacteriol* **179**: 7081–7088.

Bae, W., Xia, B., Inouye, M., and Severinov, K. (2000) *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc Natl Acad Sci USA* **97**: 7784–7789.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., *et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**: W202–8.

Baker, C.S., Morozov, I., Suzuki, K., Romeo, T., and Babitzke, P. (2002) CsrA regulates glycogen biosynthesis by preventing translation of *glgC* in *Escherichia coli*. *Mol Microbiol* **44**: 1599–1610.

Basu, A., and Yap, M.-N.F. (2016) Ribosome hibernation factor promotes *Staphylococcal* survival and differentially represses translation. *Nucleic Acids Res* **44**: 4881–4893.

Beisel, C.L., Updegrove, T.B., Janson, B.J., and Storz, G. (2012) Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J* **31**: 1961–1974.

Benhalevy, D., Bochkareva, E.S., Biran, I., and Bibi, E. (2015) Model Uracil-Rich RNAs and Membrane Protein mRNAs Interact Specifically with Cold Shock Proteins in *Escherichia coli*. *PLoS ONE* **10**: e0134413–18.

Boni, I.V., Artamonova, V.S., Tzareva, N.V., and Dreyfus, M. (2001) Non-canonical mechanism for translational control in bacteria: Synthesis of ribosomal protein S1. *EMBO J* **20**: 4222–4232.

Bonnin, R.A., and Bouloc, P. (2015) RNA Degradation in *Staphylococcus aureus*: Diversity of Ribonucleases and Their Impact. *Int J Genomics* **2015**: 395753.

Bronesky, D., Wu, Z., Marzi, S., Walter, P., Geissmann, T., Moreau, K., Vandenesch, F., Caldelari, I. and Romby, P. (2016) *Staphylococcus aureus* RNAIII and Its Regulon Link Quorum Sensing, Stress Responses, Metabolic Adaptation, and Regulation of Virulence Gene Expression. *Annu. Rev. Microbiol.*, **70**, 299–316.

Bronsard, J., Pascreau, G., Sassi, M., Mauro, T., Augagneur, Y., and Felden, B. (2017) sRNA and cis-antisense sRNA identification in Staphylococcus aureus highlights an unusual sRNA gene cluster with one encoding a secreted peptide. *Sci Rep* **7**: 603–17.

Bycroft, M., Hubbard, T.J.P., Proctor, M., Freund, S.M.V., and Murzin, A.G. (1997) The solution structure of the S1 RNA binding domain: A member of an ancient nucleic acid-binding fold. *Cell* **88**: 235–242.

Byrgazov, K., Manoharadas, S., Kaberdina, A.C., Vesper, O., and Moll, I. (2012) Direct interaction of the N-terminal domain of ribosomal protein S1 with protein S2 in Escherichia coli. *PLoS ONE* **7**: e32702.

Carpousis, A.J. (2007) *The RNA degradosome of Escherichia coli: An mRNA-degrading machine assembled on RNase E*.

Chao, Y., Li, L., Girodat, D., Förstner, K.U., Said, N., Corcoran, C., *et al.* (2017) In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol Cell* **65**: 39–51.

Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., and Vogel, J. (2012) An atlas of Hfq-bound transcripts reveals 3′ UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J* **31**: 4005–4019.

Charpentier, E., Anton, A.I., Barry, P., Alfonso, B., Fang, Y., and Novick, R.P. (2004) Novel cassette-based shuttle vector system for gram-positive bacteria. *Appl Environ Microbiol* **70**: 6076–6085.

Cheng, Z.-F., and Deutscher, M.P. (2005) An important role for RNase R in mRNA decay. *Mol Cell* **17**: 313–318.

Cherkasov, A., Hsing, M., Zoraghi, R., Foster, L.J., See, R.H., Stoynov, N., *et al.* (2011) Mapping the Protein Interaction Network in Methicillin-Resistant Staphylococcus aureus. *J Proteome Res* **10**: 1139–1150.

Cho, K.H. (2017) The structure and function of the gram-positive bacterial RNA degradosome. *Front Microbiol* **8**: 154.

Coburn, G.A., and Mackie, G.A. (1998) Reconstitution of the degradation of the mRNA for ribosomal protein S20 with purified enzymes. *Journal of Molecular Biology* **279**: 1061–1074.

Cordwell, S.J., Larsen, M.R., Cole, R.T., and Walsh, B.J. (2002) Comparative proteomics of Staphylococcus aureus and the response of methicillin-resistant and methicillin-sensitive strains to Triton X-100. *Microbiology* **148**: 2765–2781.

Court, D.L., Gan, J., Liang, Y.-H., Shaw, G.X., Tropea, J.E., Costantino, N., *et al.* (2013) *RNase III: Genetics and function; structure and mechanism.*

Cramton, S.E., Gerke, C., Schnell, N.F., Nichols, W.W., and Götz, F. (1999) The intercellular adhesion (*ica*) locus is present in *Staphylococcus aureus* and is required for biofilm formation. *Infect Immun* **67**: 5427–5433.

Crick, F. (1970) Central dogma of molecular biology. *Nature* **227**: 561–563.

Culver, G.M., Cate, J.H., Yusupova, G.Z., Yusupov, M.M., and Noller, H.F. (1999) Identification of an RNA-protein bridge spanning the ribosomal subunit interface. *Science* **285**: 2133–2135.

Czapski, T.R., and Trun, N. (2014) Expression of csp genes in *E. coli* K-12 in defined rich and defined minimal media during normal growth, and after cold-shock. *Gene* **547**: 91–97.

Dambach, M., Irnov, I., and Winkler, W.C. (2013) Association of RNAs with *Bacillus subtilis* Hfq. *PLoS ONE* **8**: e55156–17.

Davis, A.R., Gohara, D.W., and Yap, M.-N.F. (2014) Sequence selectivity of macrolide-induced translational attenuation. *Proc Natl Acad Sci USA* **111**: 15379–15384.

Delvillani, F., Papiani, G., Dehò, G., and Briani, F. (2011) S1 ribosomal protein and the interplay between translation and mRNA decay. *Nucleic Acids Res* **39**: 7702–7715.

Derman, Y., Söderholm, H., Lindström, M., and Korkeala, H. (2015) Role of *csp* genes in NaCl, pH, and ethanol stress response and motility in *Clostridium botulinum* ATCC 3502. *Food Microbiol* **46**: 463–470.

Donovan, W.P., and Kushner, S.R. (1986) Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in Escherichia coli K-12. *Proc Natl Acad Sci USA* **83**: 120–124.

Dubey, A.K., Baker, C.S., Romeo, T., and Babitzke, P. (2005) RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA* **11**: 1579–1587.

Dubey, A.K., Baker, C.S., Suzuki, K., Jones, A.D., Pandit, P., Romeo, T., and Babitzke, P. (2003) CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript. *J Bacteriol* **185**: 4450–4460.

Duval, B.D., Mathew, A., Satola, S.W., and Shafer, W.M. (2010) Altered growth, pigmentation, and antimicrobial susceptibility properties of *Staphylococcus aureus* due to loss of the major cold shock gene *cspB*. *Antimicrob Agents Chemother* **54**: 2283–2290.

Duval, M., Korepanov, A., Fuchsbauer, O., Fechter, P., Haller, A., Fabbretti, A., *et al.* (2013) Escherichia coli Ribosomal Protein S1 Unfolds Structured mRNAs Onto the Ribosome for Active Translation Initiation. *PLoS Biol* **11**: e1001731.

Endo, G., and Silver, S. (1995) CadC, the transcriptional regulatory protein of the cadmium resistance system of *Staphylococcus aureus* plasmid pI258. *J Bacteriol* **177**: 4437–4441.

Even, S., Pellegrini, O., Zig, L., Labas, V., Vinh, J., Bréchemmier-Baey, D., and Putzer, H. (2005) Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E. coli* RNase E. *Nucleic Acids Res* **33**: 2141–2152.

Fechter, P., Caldelari, I., Lioliou, E., and Romby, P. (2014) Novel aspects of RNA regulation in *Staphylococcus aureus*. *FEBS Letters* **588**: 2523–2529.

Franze de Fernandez, M.T., Eoyang, L., and August, J.T. (1968) Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219**: 588–590.

Freese, N.H., Norris, D.C., and Loraine, A.E. (2016) Integrated genome browser: Visual analytics platform for genomics. *Bioinformatics* **32**: 2089–2095.

Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2006) Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* **124**: 355–366.

Gaupp, R., Ledala, N., and Somerville, G.A. (2012) Staphylococcal response to oxidative stress. *Front Cell Infect Microbiol* **2**: 33.

Giuliodori, A.M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., *et al.* (2010) The *cspA* mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. *Mol Cell* **37**: 21–33.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS Letters* **582**: 1977–1986.

Graumann, P., Wendrich, T.M., Weber, M.H., Schröder, K., and Marahiel, M.A. (1997) A family of cold shock proteins in Bacillus subtilis is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. *Mol Microbiol* **25**: 741–756.

Graumann, P.L., and Marahiel, M.A. (1998) A superfamily of proteins that contain the cold-shock domain. *Trends Biochem Sci* **23**: 286–290.

Graumann, P.L., and Marahiel, M.A. (1999) Cold shock proteins CspB and CspC are major stationary-phase-induced proteins in *Bacillus subtilis*. *Arch Microbiol* **171**: 135–138.

Griaznova, O., and Traut, R.R. (2000) Deletion of C-terminal residues of Escherichia coli ribosomal protein L10 causes the loss of binding of one L7/L12 dimer: Ribosomes with one L7/L12 dimer are active. *Biochemistry* **39**: 4075–4081.

Gutiérrez, P., Li, Y., Osborne, M.J., Pomerantseva, E., Liu, Q., and Gehring, K. (2005) Solution structure of the carbon storage regulator protein CsrA from Escherichia coli. *J Bacteriol* **187**: 3496–3501.

Hammarlöf, D.L., Bergman, J.M., Garmendia, E., and Hughes, D. (2015) Turnover of mRNAs is one of the essential functions of RNase E. *Mol Microbiol* **98**: 34–45.

Hillier, B.J., Rodriguez, H.M., and Gregoret, L.M. (1998) Coupling protein stability and protein function in *Escherichia coli* CspA. *Fold Des* **3**: 87–93.

Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., *et al.* (2016) Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking *in vivo*. *EMBO J* **35**: 991–1011.

Iandolo, J.J., Worrell, V., Groicher, K.H., Qian, Y., Tian, R., Kenton, S., *et al.* (2002) Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of Staphylococcus aureus 8325. *Gene* **289**: 109–118.

Iben, J.R., and Draper, D.E. (2008) Specific interactions of the L10(L12)4 ribosomal protein complex with mRNA, rRNA, and L11. *Biochemistry* **47**: 2721–2731.

Ishikawa, H., Otaka, H., Maki, K., Morita, T., and Aiba, H. (2012) The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3′ poly(U) tail. *RNA* **18**: 1062–1074.

Jain, R., Devine, T., George, A.D., Chittur, S.V., Baroni, T.E., Penalva, L.O., and Tenenbaum, S.A. (2010) RIP-Chip Analysis: RNA-Binding Protein Immunoprecipitation-Microarray (Chip) Profiling. In *RNA*. Nielsen, H. (ed.). Humana Press, Totowa, NJ. pp. 247–263.

Ji, H., Jiang, H., Ma, W., and Wong, W.H. (2011) Using CisGenome to analyze ChIP-chip and ChIP-seq data. *Curr Protoc Bioinformatics* **Chapter 2**: Unit2.13–45.

Jinks-Robertson, S., and Nomura, M. (1982) Ribosomal protein S4 acts in trans as a translational repressor to regulate expression of the alpha operon in *Escherichia coli*. *J Bacteriol* **151**: 193–202.

Katzif, S., Danavall, D., Bowers, S., Balthazar, J.T., and Shafer, W.M. (2003) The major cold shock gene, *cspA*, is involved in the susceptibility of *Staphylococcus aureus* to an antimicrobial peptide of human cathepsin G. *Infect Immun* **71**: 4304–4312.

Katzif, S., Lee, E.H., Law, A.B., Tzeng, Y.L., and Shafer, W.M. (2005) CspA Regulates Pigment Production in *Staphylococcus aureus* through a SigB-Dependent Mechanism. *J Bacteriol* **187**: 8181–8184.

Kenan, D.J., Query, C.C., and Keene, J.D. (1991) RNA recognition: towards identifying determinants of specificity. *Trends Biochem Sci* **16**: 214–220.

Kim, H., Abeysirigunawarden, S.C., Chen, K., Mayerle, M., Ragunathan, K., Luthey-Schulten, Z., *et al.* (2014) Protein-guided RNA dynamics during early ribosome assembly. *Nature* **506**: 334–338.

Knox, J., Uhlemann, A.C., and Lowy, F.D. (2015) Staphylococcus aureus infections: Transmission within households and the community. *Trends in Microbiology* **23**: 437–444.

Koch, G., Yepes, A., Förstner, K.U., Wermser, C., Stengel, S.T., Modamio, J., *et al.* (2014) Evolution of Resistance to a Last-Resort Antibiotic in Staphylococcus aureus via Bacterial Competition. *Cell* **158**: 1060–1071.

Kortmann, J., and Narberhaus, F. (2012) Bacterial RNA thermometers: molecular zippers and switches. *Nat Rev Micro* **10**: 255–265.

Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**: 255–258.

Kuroda, M., Ohta, T., and Hayashi, H. (1995) Isolation and the gene cloning of an alkaline shock protein in methicillin resistant *Staphylococcus aureus*. *Biochem Biophys Res Commun* **207**: 978–984.

Lamontagne, B., and Elela, S.A. (2004) Evaluation of the RNA Determinants for Bacterial and Yeast RNase III Binding and Cleavage. *J Biol Chem* **279**: 2231–2241.

Lan, L., Cheng, A., Dunman, P.M., Missiakas, D., and He, C. (2010) Golden pigment production and virulence gene expression are affected by metabolisms in Staphylococcus aureus. *J Bacteriol* **192**: 3068–3077.

Lasa, I., and Villanueva, M. (2014) Overlapping transcription and bacterial RNA removal. *Proc Natl Acad Sci USA* **111**: 2868–2869.

Lasa, I., Toledo-Arana, A., and Gingeras, T.R. (2012) An effort to make sense of antisense transcription in bacteria. *RNA Biol* **9**: 1039–1044.

Lasa, I., Toledo-Arana, A., Dobin, A., Villanueva, M., de los Mozos, I.R., Vergara-Irigaray, M., *et al.* (2011) Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Natl Acad Sci USA* **108**: 20172–20177.

Lee, J., Jeong, K.-W., Jin, B., Ryu, K.-S., Kim, E.-H., Ahn, J.-H., and Kim, Y. (2013) Structural and dynamic features of cold-shock proteins of *Listeria monocytogenes*, a psychrophilic bacterium. *Biochemistry* **52**: 2492–2504.

Lee, J.C. (1995) Electrotransformation of Staphylococci. *Methods Mol Biol* **47**: 209–216.

Lehnik-Habrink, M., Schaffer, M., Mäder, U., Diethmaier, C., Herzberg, C., and Stülke, J. (2011) RNA processing in *Bacillus subtilis*: Identification of targets of the essential RNase Y. *Mol Microbiol* **81**: 1459–1473.

Letunic, I., Doerks, T., and Bork, P. (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* **43**: D257–D260.

Link, T.M., Valentin-Hansen, P., and Brennan, R.G. (2009) Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proceedings of the National Academy of Sciences* **106**: 19292–19297.

Lioliou, E., Sharma, C.M., Caldelari, I., Helfer, A.C., Fechter, P., Vandenesch, F., *et al.* (2012) Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* **8**: e1002782.

Liu, C.-I., Liu, G.Y., Song, Y., Yin, F., Hensler, M.E., Jeng, W.-Y., *et al.* (2008) A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. *Science* **319**: 1391–1394.

Liu, G.Y., Essex, A., Buchanan, J.T., Datta, V., Hoffman, H.M., Bastian, J.F., *et al.* (2005) Staphylococcus aureus golden pigment impairs neutrophil killing and promotes virulence through its antioxidant activity. *Journal of Experimental Medicine* **202**: 209–215.

Liu, M.Y., Gui, G., Wei, B., Preston, J.F., Oakford, L., Yüksel, Ü., *et al.* (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli. *J Biol Chem* **272**: 17502–17510.

Loepfe, C., Raimann, E., Stephan, R., and Tasara, T. (2010) Reduced host cell invasiveness and oxidative stress tolerance in double and triple csp gene family deletion mutants of *Listeria monocytogenes*. *Foodborne Pathog Dis* **7**: 775–783.

Lopez, M.M., Yutani, K., and Makhatadze, G.I. (1999) Interactions of the major cold shock protein of Bacillus subtilis CspB with single-stranded DNA templates of different base composition. *J Biol Chem* **274**: 33601–33608.

Lopez, M.M., Yutani, K., and Makhatadze, G.I. (2001) Interactions of the cold shock protein CspB from *Bacillus subtilis* with single-stranded DNA. Importance of the T base content and position within the template. *J Biol Chem* **276**: 15511–15518.

Lorenz, C., Gesell, T., Zimmermann, B., Schoeberl, U., Bilusic, I., Rajkowitsch, L., *et al.* (2010) Genomic SELEX for Hfq-binding RNAs identifies genomic aptamers predominantly in antisense transcripts. *Nucleic Acids Res* **38**: 3794–3808.

Lorenz, U., Ohlsen, K., Karch, H., Hecker, M., Thiede, A., and Hacker, J. (2000) Human antibody response during sepsis against targets expressed by methicillin resistant *Staphylococcus aureus*. *FEMS Immunol Med Microbiol* **29**: 145–153.

López-Garrido, J., Puerta-Fernández, E., and Casadesús, J. (2014) A eukaryotic-like 3' untranslated region in Salmonella enterica hilD mRNA. *Nucleic Acids Res* **42**: 5894–5906.

Lybecker, M., Zimmermann, B., Bilusic, I., Tukhtubaeva, N., and Schroeder, R. (2014) The double-stranded transcriptome of *Escherichia coli*. *Proceedings of the National Academy of Sciences* **111**: 3134–3139.

Mackie, G.A. (1998) Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**: 720–723.

Maira-Litran, T., Kropec, A., Goldmann, D.A., and Pier, G.B. (2005) Comparative Opsonic and Protective Activities of Staphylococcus aureus Conjugate Vaccines Containing Native or Deacetylated Staphylococcal Poly-N-Acetyl- -(1-6)-Glucosamine. *Infect Immun* **73**: 6752–6762.

Mathy, N., Bénard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007) 5'-to-3‘ exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5’ stability of mRNA. *Cell* **129**: 681–692.

Mathy, N., Hébert, A., Mervelet, P., Bénard, L., Dorléans, A., Li De La Sierra-Gallay, I., *et al.* (2010) *Bacillus subtilis* ribonucleases J1 and J2 form a complex with altered enzyme behaviour. *Mol Microbiol* **75**: 489–498.

Matsunaga, J., Simons, E.L., and Simons, R.W. (1996) RNase III autoregulation: Structure and function of *rncO*, the posttranscriptional “operator.” *RNA* **2**: 1228–1240.

Mattheakis, L.C., and Nomura, M. (1988) Feedback regulation of the spc operon in Escherichia coli: translational coupling and mRNA processing. *J Bacteriol* **170**: 4484–4492.

Max, K.E.A., Zeeb, M., Bienert, R., Balbach, J., and Heinemann, U. (2007) Common mode of DNA binding to cold shock domains. Crystal structure of hexathymidine bound to the domain-swapped form of a major cold shock protein from *Bacillus caldolyticus*. *FEBS J* **274**: 1265–1279.

Mäder, U., Zig, L., Kretschmer, J., Homuth, G., and Putzer, H. (2008) mRNA processing by RNases J1 and J2 affects Bacillus subtilis gene expression on a global scale. *Mol Microbiol* **70**: 183–196.

McDowall, K.J., Lin-Chaol, S., and Cohen, S.N. (1994) A + U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J Biol Chem* **269**: 10790–10796.

Mercante, J., Suzuki, K., Cheng, X., Babitzke, P., and Romeo, T. (2006) Comprehensive alanine-scanning mutagenesis of Escherichia coli CsrA defines two subdomains of critical functional importance. *J Biol Chem* **281**: 31832–31842.

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., and Thomas, P.D. (2017) PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* **45**: D183–D189.

Michaux, C., Holmqvist, E., Vasicek, E., Sharan, M., Barquist, L., Westermann, A.J., *et al.* (2017) RNA target profiles direct the discovery of virulence functions for the cold-shock proteins CspC and CspE. *Proc Natl Acad Sci USA* **65**: 201620772–48.

Michaux, C., Martini, C., Shioya, K., Ahmed Lecheheb, S., Budin-Verneuil, A., Cosette, P., *et al.* (2012) CspR, a Cold Shock RNA-Binding Protein Involved in the Long-Term Survival and the Virulence of *Enterococcus faecalis*. *J Bacteriol* **194**: 6900–6908.

Mohanty, B.K., and Kushner, S.R. (2000) Polynucleotide phosphorylase functions both as a 3“ right-arrow 5” exonuclease and a poly(A) polymerase in Escherichiacoli. *Proceedings of the National Academy of Sciences* **97**: 11966–11971.

Mohanty, B.K., and Kushner, S.R. (2003) Genomic analysis in *Escherichia coli* demonstrates differential roles for polynucleotide phosphorylase and RNase II in mRNA abundance and decay. *Mol Microbiol* **50**: 645–658.

Mohanty, B.K., and Kushner, S.R. (2011) Bacterial/archaeal/organellar polyadenylation. *WIREs RNA* **2**: 256–276.

Nakashima, K., Kanamaru, K., Mizuno, T., and Horikoshi, K. (1996) A novel member of the *cspA* family of genes that is induced by cold shock in *Escherichia coli*. *J Bacteriol* **178**: 2994–2997.

Nashimoto, H., and Nomura, M. (1970) Structure and function of bacterial ribosomes. XI. Dependence of 50S ribosomal assembly on simultaneous assembly of 30S subunits. *Proc Natl Acad Sci USA* **67**: 1440–1447.

Nevskaya, N., Tishchenko, S., Gabdoulkhakov, A., Nikonova, E., Nikonov, O., Nikulin, A., *et al.* (2005) Ribosomal protein L1 recognizes the same specific structural motif in its target sites on the autoregulatory mRNA and 23S rRNA. *Nucleic Acids Res* **33**: 478–485.

Newkirk, K., Feng, W., Jiang, W., Tejero, R., Emerson, S.D., Inouye, M., and Montelione, G.T. (1994) Solution NMR structure of the major cold shock protein (CspA) from *Escherichia coli*: identification of a binding epitope for DNA. *Proc Natl Acad Sci USA* **91**: 5114–5118.

Nomura, M. (1970) Bacterial ribosome. *Bacteriol Rev* **34**: 228–277.

Nossal, N.G., and Singer, M.F. (1968) The processive degradation of individual polyribonucleotide chains. I. Escherichia coli ribonuclease II. *J Biol Chem* **243**: 913–922.

Novick, R.P., Ross, H.F., Projan, S.J., Kornblum, J., Kreiswirth, B., and Moghazeh, S. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* **12**: 3967–3975.

Otaka, H., Ishikawa, H., Morita, T., and Aiba, H. (2011) PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc Natl Acad Sci USA* **108**: 13059–13064.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.-Y., Cohoon, M., *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* **33**: 5691–5702.

Ow, M.C., and Kushner, S.R. (2002) Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. *Genes & Development* **16**: 1102–1115.

Panja, S., and Woodson, S.A. (2012) Hfq proximity and orientation controls RNA annealing. *Nucleic Acids Res* **40**: 8690–8697.

Panja, S., Schu, D.J., and Woodson, S.A. (2013) Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Res* **41**: 7536–7546.

Park, H., McGibbon, L.C., Potts, A.H., Yakhnin, H., Romeo, T., and Babitzke, P. (2017) Translational repression of the RpoS antiadapter IraD by CsrA is mediated via translational coupling to a short upstream open reading frame. *mBio* **8**: e01355–17.

Pelz, A., Wieland, K.-P., Putzbach, K., Hentschel, P., Albert, K., and Götz, F. (2005) Structure and biosynthesis of staphyloxanthin from *Staphylococcus aureus*. *J Biol Chem* **280**: 32493–32498.

Peng, Y., Curtis, J.E., Fang, X., and Woodson, S.A. (2014) Structural model of an mRNA in complex with the bacterial chaperone Hfq. *Proceedings of the National Academy of Sciences* **111**: 17134–17139.

Peng, Y., Soper, T.J., and Woodson, S.A. (2014) Positional effects of AAN motifs in rpoS regulation by sRNAs and Hfq. *Journal of Molecular Biology* **426**: 275–285.

Pertzev, A.V. (2006) Characterization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of *Escherichia coli* ribonuclease III. *Nucleic Acids Res* **34**: 3708–3721.

Phadtare, S. (2004) Recent developments in bacterial cold-shock response. *Curr Issues Mol Biol* **6**: 125–136.

Phadtare, S., and Inouye, M. (1999) Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of *Escherichia coli*. *Mol Microbiol* **33**: 1004–1014.

Phadtare, S., and Severinov, K. (2005) Nucleic acid melting by *Escherichia coli* CspE. *Nucleic Acids Res* **33**: 5583–5590.

Phadtare, S., and Severinov, K. (2010) RNA remodeling and gene regulation by cold shock proteins. *RNA Biol* **7**: 788–795.

Phadtare, S., Inouye, M., and Severinov, K. (2002) The nucleic acid melting activity of Escherichia coli CspE is critical for transcription antitermination and cold acclimation of cells. *J Biol Chem* **277**: 7239–7245.

Phadtare, S., Severinov, K., and Inouye, M. (2003) Assay of transcription antitermination by proteins of the CspA family. *Meth Enzymol* **371**: 460–471.

Phadtare, S., Tyagi, S., Inouye, M., and Severinov, K. (2002) Three amino acids in *Escherichia coli* CspE surface-exposed aromatic patch are critical for nucleic acid melting activity leading to transcription antitermination and cold acclimation of cells. *J Biol Chem* **277**: 46706–46711.

Philippe, C., Eyermann, F., Benard, L., Portier, C., Ehresmann, B., and Ehresmann, C. (1993) Ribosomal protein S15 from *Escherichia coli* modulates its own translation by trapping the ribosome on the mRNA initiation loading site. *Proc Natl Acad Sci USA* **90**: 4394–4398.

Pop, C., Rouskin, S., Ingolia, N.T., Han, L., Phizicky, E.M., Weissman, J.S., and Koller, D. (2014) Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. *Mol Syst Biol* **10**: 770–770.

Potts, A.H., Vakulskas, C.A., Pannuri, A., Yakhnin, H., Babitzke, P., and Romeo, T. (2017) Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. *Nature Communications* **8**: 1596.

Qu, X., Lancaster, L., Noller, H.F., Bustamante, C., and Tinoco, I. (2012) Ribosomal protein S1 unwinds double-stranded RNA in multiple steps. *Proc Natl Acad Sci USA* **109**: 14458–14463.

Raibaud, S., Vachette, P., Guillier, M., Allemand, F., Chiaruttini, C., and Dardel, F. (2003) How bacterial ribosomal protein L20 assembles with 23 S ribosomal RNA and its own messenger RNA. *J Biol Chem* **278**: 36522–36530.

Redder, P., Hausmann, S., Khemici, V., Yasrebi, H., and Linder, P. (2015) Bacterial versatility requires DEAD-box RNA helicases. *FEMS Microbiology Reviews* **39**: 392–412.

Ren, G.-X., Guo, X.-P., and Sun, Y.-C. (2017) Regulatory 3′ Untranslated Regions of Bacterial mRNAs. *Front Microbiol* **8**: 633–6.

Richards, J., Liu, Q., Pellegrini, O., Celesnik, H., Yao, S., Bechhofer, D.H., *et al.* (2011) An RNA Pyrophosphohydrolase Triggers 5. *Mol Cell* **43**: 940–949.

Robert, F., and Brakier-Gingras, L. (2001) Ribosomal protein S7 from Escherichia coli uses the same determinants to bind 16S ribosomal RNA and its messenger RNA. *Nucleic Acids Res* **29**: 677–682.

Romeo, T., Vakulskas, C.A., and Babitzke, P. (2013) Post-transcriptional regulation on a global scale: Form and function of Csr/Rsm systems. *Environmental Microbiology* **15**: 313–324.

Ruiz de Los Mozos, I., Vergara-Irigaray, M., Segura, V., Villanueva, M., Bitarte, N., Saramago, M., *et al.* (2013) Base pairing interaction between 5“- and 3-”UTRs controls icaR mRNA translation in *Staphylococcus aureus*. *PLoS Genet* **9**: e1004001.

Sabnis, N.A., Yang, H., and Romeo, T. (1995) Pleiotropic regulation of central carbohydrate metabolism in *Escherichia coli* via the gene *csrA*. *J Biol Chem* **270**: 29096–29104.

Sachs, R., Max, K.E.A., Heinemann, U., and Balbach, J. (2012) RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution. *RNA* **18**: 65–76.

Sahukhal, G.S., and Elasri, M.O. (2014) Identification and characterization of an operon, *msaABCR*, that controls virulence and biofilm development in *Staphylococcus aureus*. *BMC Microbiol* **14**: 154.

Saito, K., and Nomura, M. (1994) Post-transcriptional regulation of the str operon in *Escherichia coli*: Structural and mutational analysis of the target site for translational repressor S7. *Journal of Molecular Biology* **235**: 125–139.

Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C., and Bontems, F. (2009) Probing the relationship between gram-negative and gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res* **37**: 5578–5588.

Sauer, E., and Weichenrieder, O. (2011) Structural basis for RNA 3′ -end recognition by Hfq. *Proc Natl Acad Sci USA* **108**: 13065–13070.

Sauer, E., Schmidt, S., and Weichenrieder, O. (2012) Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc Natl Acad Sci USA* **109**: 9396–9401.

Sauter, C., Basquin, J., and Suck, D. (2003) Sm-like proteins in Eubacteria: The crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res* **31**: 4091–4098.

Schindelin, H., Jiang, W., Inouye, M., and Heinemann, U. (1994) Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc Natl Acad Sci USA* **91**: 5119–5123.

Schmid, B., Klumpp, J., Raimann, E., Loessner, M.J., Stephan, R., and Tasara, T. (2009) Role of cold shock proteins in growth of *Listeria monocytogenes* under cold and osmotic stress conditions. *Appl Environ Microbiol* **75**: 1621–1627.

Schröder, K., Graumann, P., Schnuchel, A., Holak, T.A., and Marahiel, M.A. (1995) Mutational analysis of the putative nucleic acid-binding surface of the cold-shock domain, CspB, revealed an essential role of aromatic and basic residues in binding of single-stranded DNA containing the Y-box motif. *Mol Microbiol* **16**: 699–708.

Schu, D.J., Zhang, A., Gottesman, S., and Storz, G. (2015) Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *EMBO J* **34**: 2557–2573.

Schumacher, M.A., Pearson, R.F., Møller, T., Valentin-Hansen, P., and Brennan, R.G. (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: A bacterial Sm-like protein. *EMBO J* **21**: 3546–3556.

Segura, V., Toledo-Arana, A., Uzqueda, M., Lasa, I., and Muñoz-Barrutia, A. (2012) Wavelet-based detection of transcriptional activity on a novel *Staphylococcus aureus* tiling microarray. *BMC Bioinformatics* **13**: 222.

Sengupta, J., Agrawal, R.K., and Frank, J. (2001) Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc Natl Acad Sci USA* **98**: 11991–11996.

Serganov, A., and Nudler, E. (2013) A decade of riboswitches. *Cell* **152**: 17–24.

Serganov, A., Polonskaia, A., Ehresmann, B., Ehresmann, C., and Patel, D.J. (2003) Ribosomal protein S15 represses its own translation via adaptation of an rRNA-like fold within its mRNA. *EMBO J* **22**: 1898–1908.

Shahbabian, K., Jamalli, A., Zig, L., and Putzer, H. (2009) RNase Y, a novel endoribonuclease, initiates riboswitch turnover in Bacillus subtilis. *EMBO J* **28**: 3523–3533.

Shenhar, Y., Biran, D., and Ron, E.Z. (2012) Resistance to environmental stress requires the RNA chaperones CspC and CspE. *Environmental Microbiology Reports* **4**: 532–539.

Shilov, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., *et al.* (2007) The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics* **6**: 1638–1655.

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., *et al.* (2008) Deep Sequencing Analysis of Small Noncoding RNA and mRNA Targets of the Global Post-Transcriptional Regulator, Hfq. *PLoS Genet* **4**: e1000163.

Skinner, M.E., Uzilov, A.V., Stein, L.D., Mungall, C.J., and Holmes, I.H. (2009) JBrowse: a next-generation genome browser. *Genome Res* **19**: 1630–1638.

Skouv, J., Schnier, J., Rasmussen, M.D., Subramanian, A.R., and Pedersen, S. (1990) Ribosomal protein S1 of *Escherichia coli* is the effector for the regulation of its own synthesis. *J Biol Chem* **265**: 17044–17049.

Smirnov, A., Förstner, K.U., Holmqvist, E., Otto, A., Günster, R., Becher, D., *et al.* (2016) Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc Natl Acad Sci USA* **113**: 11591–11596.

Soper, T.J., and Woodson, S.A. (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* **14**: 1907–1917.

Sorensen, M.A., Fricke, J., and Pedersen, S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in Escherichia coli in vivo. *Journal of Molecular Biology* **280**: 561–569.

Spickler, C., and Mackie, G.A. (2000) Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J Bacteriol* **182**: 2422–2427.

Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Castillo, L.P., Ray, D., *et al.* (2011) Analysis of *Escherichia coli* RNase E and RNase III activity in vivo using tiling microarrays. *Nucleic Acids Res* **39**: 3188–3203.

Stelzl, U., Zengel, J.M., Tovbina, M., Walker, M., Nierhaus, K.H., Lindahl, L., and Patel, D.J. (2003) RNA-structural mimicry in Escherichia coli ribosomal protein L4-dependent regulation of the S10 operon. *J Biol Chem* **278**: 28237–28245.

Storz, G., Vogel, J., and Wassarman, K.M. (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* **43**: 880–891.

Takyar, S., Hickerson, R.P., and Noller, H.F. (2005) mRNA helicase activity of the ribosome. *Cell* **120**: 49–58.

Tanaka, T., Mega, R., Kim, K., Shinkai, A., Masui, R., Kuramitsu, S., and Nakagawa, N. (2012) A non-cold-inducible cold shock protein homolog mainly contributes to translational control under optimal growth conditions. *FEBS Journal* **279**: 1014–1029.

Tang, K.L., Caffrey, N.P., Nobrega, D.B., Cork, S.C., Ronksley, P.E., Barkema, H.W., *et al.* (2017) Articles Restricting the use of antibiotics in food-producing animals and its associations with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis. *The Lancet Planetary Health* **1**: e316–e327.

Thomas, M.S., Bedwell, D.M., and Nomura, M. (1987) Regulation of alpha operon gene expression in *Escherichia coli*. A novel form of translational coupling. *Journal of Molecular Biology* **196**: 333–345.

Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., *et al.* (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**: 950–956.

Tomasini, A., Francois, P., Howden, B.P., Fechter, P., Romby, P., and Caldelari, I. (2014) The importance of regulatory RNAs in Staphylococcus aureus. *Infect Genet Evol* **21**: 616–626.

Tong, S.Y.C., Davis, J.S., Eichenberger, E., Holland, T.L., and Fowler, V.G. (2015) *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* **28**: 603–661.

Torres, M., Condon, C., Balada, J.M., Squires, C., and Squires, C.L. (2001) Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J* **20**: 3811–3820.

Updegrove, T.B., Zhang, A., and Storz, G. (2016) Hfq: The flexible RNA matchmaker. *Current Opinion in Microbiology* **30**: 133–138.

Uppalapati, C.K., Gutierrez, K.D., Buss-Valley, G., and Katzif, S. (2017) Growth-dependent activity of the cold shock cspA promoter + 5′ UTR and production of the protein CspA in *Staphylococcus aureus* Newman. *BMC Res Notes* **10**: 232.

Valle, J., Toledo-Arana, A., Berasain, C., Ghigo, J.-M., Amorena, B., Penadés, J.R., and Lasa, I. (2003) SarA and not sigma B is essential for biofilm development by *Staphylococcus aureus*. *Mol Microbiol* **48**: 1075–1087.

Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M., and Frank, J. (2003) Locking and unlocking of ribosomal motions. *Cell* **114**: 123–134.

Van Assche, E., Van Puyvelde, S., Vanderleyden, J., and Steenackers, H.P. (2015) RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front Microbiol* **6**: 141.

Vanzo, N.F., Li, Y.S., Py, B., Blum, E., Higgins, C.F., Raynal, L.C., *et al.* (1998) Ribonuclease E organizes the protein interactions in the Escherichia coli RNA degradosome. *Genes & Development* **12**: 2770–2781.

Većerek, B., Moll, I., and Bläsi, U. (2005) Translational autocontrol of the *Escherichia coli hfq* RNA chaperone gene. *RNA* **11**: 976–984.

Wagner, E.G.H. (2013) Cycling of RNAs on Hfq. *RNA Biol* **10**: 619–626.

Waldminghaus, T., and Skarstad, K. (2010) ChIP on Chip: surprising results are often artifacts. *BMC Genomics* **11**: 414.

Wang, N., Yamanaka, K., and Inouye, M. (1999) CspI, the ninth member of the CspA family of *Escherichia coli*, is induced upon cold shock. *J Bacteriol* **181**: 1603–1609.

Waters, L.S., and Storz, G. (2009) Regulatory RNAs in Bacteria. *Cell* **136**: 615–628.

Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., *et al.* (2017) Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res* **45**: D535–D542.

Weilbacher, T., Suzuki, K., Dubey, A.K., Wang, X., Gudapaty, S., Morozov, I., *et al.* (2003) A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol Microbiol* **48**: 657–670.

Wieden, H.-J., Wintermeyer, W., and Rodnina, M.V. (2001) A common structural motif in elongation factor Ts and ribosomal protein L7/12 may be involved in the interaction with elongation factor Tu. *J Mol Evol* **52**: 129–136.

Willimsky, G., Bang, H., Fischer, G., and Marahiel, M.A. (1992) Characterization of *cspB*, a *Bacillus subtilis* inducible cold shock gene affecting cell viability at low temperatures. *J Bacteriol* **174**: 6326–6335.

Xia, B., Ke, H., and Inouye, M. (2001) Acquirement of cold sensitivity by quadruple deletion of the *cspA* family and its suppression by PNPase S1 domain in *Escherichia coli*. *Mol Microbiol* **40**: 179–188.

Xu, F., and Cohen, S.N. (1995) RNA degradation in Escherichia coli regulated by 3“ adenylation and 5” phosphorylation. *Nature* **374**: 180–183.

Yakhnin, A.V., Baker, C.S., Vakulskas, C.A., Yakhnin, H., Berezin, I., Romeo, T., and Babitzke, P. (2013) CsrA activates *flhDC* expression by protecting *flhDC* mRNA from RNase E-mediated cleavage. *Mol Microbiol* **87**: 851–866.

Yakhnin, H., Baker, C.S., Berezin, I., Evangelista, M.A., Rassin, A., Romeo, T., and Babitzke, P. (2011) CsrA Represses Translation of sdiA, Which Encodes the N-acylhomoserine-L-lactone receptor of Escherichia coli, by Binding exclusively within the coding region of sdiA mRNA. *J Bacteriol* **193**: 6162–6170.

Yakhnin, H., Yakhnin, A.V., Baker, C.S., Sineva, E., Berezin, I., Romeo, T., and Babitzke, P. (2011) Complex regulation of the global regulatory gene csrA: CsrA-mediated translational repression, transcription from five promoters by $E\sigma^{70}$ and $E\sigma(S)$, and indirect transcriptional activation by CsrA. *Mol Microbiol* **81**: 689–704.

Yamanaka, K., Fang, L., and Inouye, M. (1998) The CspA family in *Escherichia coli*: Multiple gene duplication for stress adaptation. *Mol Microbiol* **27**: 247–255.

Yusupova, G., Jenner, L., Rees, B., Moras, D., and Yusupov, M. (2006) Structural basis for messenger RNA movement on the ribosome. *Nature* **444**: 391–394.

Zeeb, M., Max, K.E.A., Weininger, U., Löw, C., Sticht, H., and Balbach, J. (2006) Recognition of T-rich single-stranded DNA by the cold shock protein Bs-CspB in solution. *Nucleic Acids Res* **34**: 4561–4571.

Zengel, J.M., Jerauld, A., Walker, A., Wahl, M.C., and Lindahl, L. (2003) The extended loops of ribosomal proteins L4 and L22 are not required for ribosome assembly or L4-mediated autogenous control. *RNA* **9**: 1188–1197.

Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G., and Gottesman, S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* **50**: 1111–1124.

Zhu, H., Mao, X.-J., Guo, X.-P., and Sun, Y.-C. (2015) The hmsT 3' untranslated region mediates c-di-GMP metabolism and biofilm formation in Yersinia pestis. *Mol Microbiol*.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.

ANNEXS

Annex I. Results of comparative label-free LC-MS-based proteomics. Up- and down-regulated proteins with a fold change ratio (Δ*cspA* vs WT) higher than 2 and a P-value lower than 0.05 are included. Colored rows show proteins encoded by CspA-targeted transcripts. Specifically, yellow rows highlight proteins in which the CspA-binding peak was nearby or included in the CDS. Orange rows indicate proteins encoded by polycistronic transcripts, where the CspA-binding peak was not contained in the CDS. The red row emphasizes CspA, which is deleted in the Δ*cspA* strain. The shown gene IDs correspond to *S. aureus* NCTC 8325.

Gene ID	Fold change	Anova (p)	PATRIC Description
SAOUHSC_01255	Down	0,027	Zinc protease
SAOUHSC_00615	Down	0,005	5'-nucleotidase YjjG (EC 3.1.3.5)
SAOUHSC_00702	Down	0,000	Hypothetical protein (FIG01108024)
SAOUHSC_00734	Down	0,000	Putative 5'(3')-deoxyribonucleotidase (EC 3.1.3.-)
SAOUHSC_00996	-1210,0	0,000	Hypothetical protein (FIG01108206)
SAOUHSC_02409	-122,0	0,000	Arginase (EC 3.5.3.1)
SAOUHSC_01403	-43,9	0,000	Cold shock protein CspA
SAOUHSC_02239	-32,8	0,000	Phage integrase
SAOUHSC_02355	-30,7	0,008	Hypothetical protein
SAOUHSC_01592	-24,1	0,032	Ferric uptake regulation protein FUR
SAOUHSC_02460	-23,0	0,000	oxidoreductase, aldo/keto reductase family
SAOUHSC_00051	-22,9	0,037	Phosphatidylinositol-specific phospholipase C (EC 4.6.1.13)
SAOUHSC_01420	-15,5	0,002	Response regulator ArlR
SAOUHSC_00239	-10,2	0,015	Ribokinase (EC 2.7.1.15)
SAOUHSC_00890	-9,8	0,002	Hypothetical protein (FIG01107890)
SAOUHSC_00755	-8,4	0,018	FIG01108099: hypothetical protein
SAOUHSC_02651	-8,2	0,001	acetyltransferase, GNAT family
SAOUHSC_02441	-7,8	0,000	Alkaline shock protein 23
SAOUHSC_00552	-7,7	0,006	Glucosamine-6-phosphate deaminase (EC 3.5.99.6)
SAOUHSC_01988	-7,2	0,004	tRNA (cytidine(34)-2'-O)-methyltransferase (EC 2.1.1.207)
SAOUHSC_00475	-7,1	0,030	Peptidyl-tRNA hydrolase (EC 3.1.1.29)
SAOUHSC_02899	-6,7	0,001	Hypothetical protein (FIG01108025)
SAOUHSC_01008	-6,2	0,003	Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)
SAOUHSC_02778	-6,2	0,001	Oxidoreductase, short-chain dehydrogenase/reductase family
SAOUHSC_01012	-6,2	0,007	Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit
SAOUHSC_01870	-5,8	0,016	Ribosomal small subunit pseudouridine synthase A (EC 4.2.1.70)
SAOUHSC_00556	-5,4	0,034	L-Proline/Glycine betaine transporter ProP
SAOUHSC_01158	-5,2	0,000	Cell division initiation protein DivIVA
SAOUHSC_00720	-5,2	0,000	Queuosine biosynthesis QueD, PTPS-I
SAOUHSC_01248	-5,1	0,028	tRNA pseudouridine synthase B (EC 4.2.1.70)
SAOUHSC_01174	-5,0	0,034	Hypothetical protein (FIG01107914)
SAOUHSC_00342	-4,9	0,015	Chromosome (plasmid) partitioning protein ParB
SAOUHSC_02150	-4,8	0,001	Thioredoxin
SAOUHSC_02013	-4,6	0,000	ThiJ/PfpI family protein
SAOUHSC_02403	-4,4	0,002	Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17)
SAOUHSC_02542	-4,4	0,003	Molybdopterin biosynthesis protein MoeA
SAOUHSC_01015	-4,3	0,014	Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)
SAOUHSC_02357	-4,3	0,005	Sua5 subfamily, required for N6-threonylcarbamoyl adenosine t(6)A37 modification in tRNA
SAOUHSC_01593	-4,2	0,001	ADP-ribose pyrophosphatase (EC 3.6.1.13)
SAOUHSC_00553	-4,1	0,001	D-arabino-3-hexulose 6-phosphate formaldehyde lyase
SAOUHSC_01010	-4,0	0,005	Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
SAOUHSC_02976	-3,9	0,034	Mannose-6-phosphate isomerase (EC 5.3.1.8)
SAOUHSC_02900	-3,7	0,000	Predicted hydrolase/acyltransferase
SAOUHSC_01456	-3,7	0,009	Uncharacterized iron-regulated membrane protein; Iron-uptake factor PiuB
SAOUHSC_02536	-3,6	0,030	Molybdenum cofactor biosynthesis protein MoaA
SAOUHSC_01009	-3,6	0,008	Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
SAOUHSC_02860	-3,5	0,000	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)
SAOUHSC_02366	-3,5	0,000	Fructose-bisphosphate aldolase class II (EC 4.1.2.13)
SAOUHSC_01139	-3,4	0,020	unknown conserved protein in B. subtilis

Continued in the following page

Annex I. Continued.

Gene ID	Fold change	Anova (p)	PATRIC Description
SAOUHSC_02434	-3,3	0,007	Siderophore synthetase superfamily, group B
SAOUHSC_00875	-3,2	0,005	NADH dehydrogenase (EC 1.6.99.3)
SAOUHSC_01198	-3,2	0,000	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)
SAOUHSC_01597	-3,2	0,013	Pyrroline-5-carboxylate reductase (EC 1.5.1.2)
SAOUHSC_02091	-3,2	0,005	Acyl-CoA hydrolase (EC 3.1.2.20)
SAOUHSC_01621	-3,1	0,004	Transcription termination protein NusB
SAOUHSC_01250	-3,1	0,010	SSU ribosomal protein S15p (S13e)
SAOUHSC_02298	-3,1	0,000	RNA polymerase sigma factor SigB
SAOUHSC_00946	-3,1	0,004	Na+/H+ antiporter
SAOUHSC_02631	-3,1	0,001	TetR family regulatory protein of MDR cluster
SAOUHSC_00019	-3,1	0,012	Adenylosuccinate synthetase (EC 6.3.4.4)
SAOUHSC_01706	-3,1	0,000	Hypothetical protein (FIG01108267)
SAOUHSC_02961	-3,0	0,017	Transcriptional regulator, MarR family
SAOUHSC_00554	-3,0	0,001	6-phospho-3-hexuloisomerase
SAOUHSC_00696	-3,0	0,016	Hypothetical protein (FIG01108122)
SAOUHSC_01452	-3,0	0,038	Alanine dehydrogenase (EC 1.4.1.1)
SAOUHSC_01858	-3,0	0,001	Phenylalanyl-tRNA synthetase domain protein (Bsu YtpR)
SAOUHSC_01786	-2,9	0,004	Translation initiation factor 3
SAOUHSC_00865	-2,9	0,001	Hypothetical NagD-like phosphatase
SAOUHSC_01606	-2,9	0,004	Peptidase T (EC 3.4.11.4)
SAOUHSC_01869	-2,9	0,000	Hypothetical protein
SAOUHSC_02774	-2,9	0,001	Hypothetical protein (FIG01108532)
SAOUHSC_02436	-2,9	0,018	Uncharacterized siderophore biosynthesis protein
SAOUHSC_01323	-2,9	0,001	Hydrolase (HAD superfamily)
SAOUHSC_01199	-2,9	0,000	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
SAOUHSC_02381	-2,8	0,000	Iron-binding ferritin-like antioxidant protein / Ferroxidase (EC 1.16.3.1)
SAOUHSC_02859	-2,8	0,000	Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)
SAOUHSC_01092	-2,8	0,047	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
SAOUHSC_02581	-2,8	0,002	Hypothetical protein similar to YrhD
SAOUHSC_01189	-2,8	0,005	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
SAOUHSC_02276	-2,8	0,000	MutS-related protein, family 1
SAOUHSC_00613	-2,8	0,000	Vitamin B12 ABC transporter, B12-binding component BtuF
SAOUHSC_02527	-2,8	0,005	tRNA-dependent lipid II--glycine ligase (FmhB)
SAOUHSC_01818	-2,7	0,014	Alanine dehydrogenase (EC 1.4.1.1)
SAOUHSC_01014	-2,7	0,009	Amidophosphoribosyltransferase (EC 2.4.2.14)
SAOUHSC_01018	-2,7	0,014	Phosphoribosylamine--glycine ligase (EC 6.3.4.13)
SAOUHSC_00644	-2,7	0,000	Teichoic acid biosynthesis protein X
SAOUHSC_00940	-2,7	0,003	Adenylate cyclase
SAOUHSC_01013	-2,7	0,022	Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)
SAOUHSC_02335	-2,7	0,031	Hypothetical protein
SAOUHSC_02501	-2,6	0,004	LSU ribosomal protein L24p (L26e)
SAOUHSC_01166	-2,6	0,008	Aspartate carbamoyltransferase (EC 2.1.3.2)
SAOUHSC_01385	-2,6	0,000	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)
SAOUHSC_02908	-2,6	0,000	Ribulosamine/erythrulosamine 3-kinase potentially involved in protein deglycation
SAOUHSC_00309	-2,6	0,002	Hypothetical protein (FIG01108282)
SAOUHSC_01017	-2,6	0,055	IMP cyclohydrolase (EC 3.5.4.10)
SAOUHSC_01007	-2,6	0,034	Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)
SAOUHSC_02754	-2,6	0,003	YbbL ABC transporter ATP-binding protein
SAOUHSC_00454	-2,6	0,024	DNA polymerase III delta prime subunit (EC 2.7.7.7)
SAOUHSC_01138	-2,6	0,000	Uncharacterized N-acetyltransferase BT9727_3663 (EC 2.3.1.-)
SAOUHSC_01773	-2,6	0,010	Uroporphyrinogen-III synthase (EC 4.2.1.75)
SAOUHSC_01365	-2,6	0,006	Deblocking aminopeptidase (EC 3.4.11.-)
SAOUHSC_00436	-2,5	0,014	Glutamate synthase [NADPH] small chain (EC 1.4.1.13)
SAOUHSC_01802	-2,5	0,029	Citrate synthase (si) (EC 2.3.3.1)
SAOUHSC_01795	-2,5	0,009	Dephospho-CoA kinase (EC 2.7.1.24)
SAOUHSC_01330	-2,5	0,011	GMP reductase (EC 1.7.1.7)
SAOUHSC_00794	-2,5	0,002	Central glycolytic genes regulator
SAOUHSC_00017	-2,5	0,006	LSU ribosomal protein L9p
SAOUHSC_00153	-2,5	0,005	Pyruvate decarboxylase (EC 4.1.1.1); Alpha-keto-acid decarboxylase (EC 4.1.1.-)

Continued in the following page

Annex I. Continued.

Gene ID	Fold change	Anova (p)	PATRIC Description
SAOUHSC_01376	-2,5	0,002	Hypothetical protein (FIG01108123)
SAOUHSC_01987	-2,5	0,000	Hypothetical protein (FIG001583), contains S4-like RNA binding domain
SAOUHSC_00574	-2,5	0,001	Phosphate acetyltransferase (EC 2.3.1.8)
SAOUHSC_01961	-2,5	0,000	Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)
SAOUHSC_01702	-2,5	0,001	5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
SAOUHSC_01749	-2,5	0,005	S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)
SAOUHSC_02566	-2,5	0,001	Transcriptional regulator SarR (Staphylococcal accessory regulator R)
SAOUHSC_01398	-2,4	0,042	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase (EC 2.3.1.89)
SAOUHSC_01695	-2,4	0,002	Hypothetical protein
SAOUHSC_00487	-2,4	0,003	Chaperonin (heat shock protein 33)
SAOUHSC_01969	-2,4	0,007	Hypothetical protein (FIG001583), contains S4-like RNA binding domain
SAOUHSC_02380	-2,4	0,004	Purine nucleoside phosphorylase (EC 2.4.2.1)
SAOUHSC_00520	-2,4	0,010	LSU ribosomal protein L10p (P0)
SAOUHSC_01810	-2,4	0,006	NADP-dependent malic enzyme (EC 1.1.1.40)
SAOUHSC_02316	-2,4	0,002	Cold-shock DEAD-box protein A
SAOUHSC_01415	-2,4	0,000	Hypothetical protein (FIG01107869)
SAOUHSC_00024	-2,4	0,011	Zn-dependent hydrolase (beta-lactamase superfamily)
SAOUHSC_00469	-2,4	0,000	Protein of unknown function identified by role in sporulation (SpoVG)
SAOUHSC_01156	-2,4	0,005	Hypothetical protein (FIG001583), contains S4-like RNA binding domain
SAOUHSC_01016	-2,4	0,028	Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)
SAOUHSC_01107	-2,4	0,003	Nucleoside 5-triphosphatase RdgB (dHAPTP, dITP, XTP-specific) (EC 3.6.1.15)
SAOUHSC_00893	-2,4	0,009	Putative NADH-dependent flavin oxidoreductase
SAOUHSC_01886	-2,4	0,010	6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.78)
SAOUHSC_00906	-2,4	0,006	Fumarylacetoacetate hydrolase family protein
SAOUHSC_02544	-2,3	0,000	Molybdenum cofactor biosynthesis protein MoaB
SAOUHSC_01214	-2,3	0,028	50S ribosomal subunit maturation GTPase RbgA (B. subtilis YlqF)
SAOUHSC_01614	-2,3	0,000	Dihydrolipoamide dehydrogenase of branched-chain alpha-keto acid dehydrogenase
SAOUHSC_00336	-2,3	0,014	3-ketoacyl-CoA thiolase (EC 2.3.1.16) @ Acetyl-CoA acetyltransferase (EC 2.3.1.9)
SAOUHSC_01314	-2,3	0,008	Two-component response regulator
SAOUHSC_01771	-2,3	0,002	Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
SAOUHSC_02447	-2,3	0,002	Putative oxidoreductase YncB
SAOUHSC_00879	-2,3	0,001	Cytosol aminopeptidase PepA (EC 3.4.11.1)
SAOUHSC_02043	-2,3	0,029	Phage head protein [SA bacteriophages 11, Mu50B] / Phage major capsid protein
SAOUHSC_00533	-2,3	0,000	chaperone protein HchA
SAOUHSC_02490	-2,3	0,001	Adenylate kinase (EC 2.7.4.3)
SAOUHSC_00920	-2,3	0,001	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII (EC 2.3.1.41)
SAOUHSC_02512	-2,3	0,017	LSU ribosomal protein L3p (L3e)
SAOUHSC_01821	-2,3	0,000	Adenine-specific methyltransferase (EC 2.1.1.72)
SAOUHSC_01977	-2,3	0,019	Hypothetical protein
SAOUHSC_02299	-2,2	0,000	Serine-protein kinase RsbW (EC 2.7.11.1)
SAOUHSC_00834	-2,2	0,015	Thioredoxin
SAOUHSC_00688	-2,2	0,000	Lysine decarboxylase family
SAOUHSC_00771	-2,2	0,021	Peptide chain release factor 2
SAOUHSC_00942	-2,2	0,001	GTP pyrophosphokinase (EC 2.7.6.5)
SAOUHSC_00797	-2,2	0,000	Triosephosphate isomerase (EC 5.3.1.1)
SAOUHSC_01201	-2,2	0,002	Acyl carrier protein / HmrB protein involved in methicillin resistance
SAOUHSC_02553	-2,2	0,008	Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)
SAOUHSC_00951	-2,2	0,001	2H phosphoesterase superfamily protein similar to Bsu1186 (yjcG)
SAOUHSC_03051	-2,2	0,004	rRNA small subunit 7-methylguanosine (m7G) methyltransferase GidB
SAOUHSC_02659	-2,2	0,018	Transcriptional regulator, TetR family
SAOUHSC_01632	-2,1	0,000	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)
SAOUHSC_00517	-2,1	0,000	Transcription antitermination protein NusG
SAOUHSC_01175	-2,1	0,018	Fibronectin/fibrinogen-binding protein
SAOUHSC_02445	-2,1	0,010	Bifunctional protein: zinc-containing alcohol dehydrogenase; quinone oxidoreductase
SAOUHSC_02604	-2,1	0,001	Dehydrogenase
SAOUHSC_01055	-2,1	0,025	Inositol-1-monophosphatase (EC 3.1.3.25)
SAOUHSC_00441	-2,1	0,000	Acetyltransferase (GNAT) family protein
SAOUHSC_01958	-2,1	0,050	Hypothetical protein

Continued in the following page

Annex I. Continued.

Gene ID	Fold change	Anova (p)	PATRIC Description
SAOUHSC_02011	-2,1	0,023	Regulatory protein RecX
SAOUHSC_01617	-2,1	0,019	Arginine pathway regulatory protein ArgR, repressor of arg regulon
SAOUHSC_00756	-2,1	0,035	Glycerate kinase (EC 2.7.1.31)
SAOUHSC_01634	-2,1	0,014	Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)
SAOUHSC_01822	-2,1	0,006	Thiol peroxidase, Tpx-type (EC 1.11.1.15)
SAOUHSC_01203	-2,1	0,000	Ribonuclease III (EC 3.1.26.3)
SAOUHSC_01845	-2,1	0,015	Formate--tetrahydrofolate ligase (EC 6.3.4.3)
SAOUHSC_01983	-2,1	0,027	Fumarate hydratase class II (EC 4.2.1.2)
SAOUHSC_01172	-2,1	0,007	Orotate phosphoribosyltransferase (EC 2.4.2.10)
SAOUHSC_02377	-2,1	0,008	Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)
SAOUHSC_01879	-2,1	0,028	Repressor of toxins Rot
SAOUHSC_02146	-2,1	0,028	Choloylglycine hydrolase (EC 3.5.1.24)
SAOUHSC_02912	-2,1	0,004	Putative DNA binding 3-demethylubiquinone-9 3-methyltransferase domain protein
SAOUHSC_02926	-2,0	0,000	Fructose-bisphosphate aldolase class I (EC 4.1.2.13)
SAOUHSC_02801	-2,0	0,000	UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)
SAOUHSC_01168	-2,0	0,021	Dihydroorotase (EC 3.5.2.3)
SAOUHSC_01653	-2,0	0,001	Manganese superoxide dismutase (EC 1.15.1.1); Superoxide dismutase [Fe] (EC 1.15.1.1)
SAOUHSC_00894	-2,0	0,000	Acetylornithine aminotransferase 2 (EC 2.6.1.11)
SAOUHSC_00244	-2,0	0,008	Hypothetical protein (FIG01108649)
SAOUHSC_01222	-2,0	0,023	DNA topoisomerase I (EC 5.99.1.2)
SAOUHSC_00796	-2,0	0,000	Phosphoglycerate kinase (EC 2.7.2.3)
SAOUHSC_01839	-2,0	0,005	Tyrosyl-tRNA synthetase (EC 6.1.1.1)
SAOUHSC_01497	-2,0	0,000	L-asparaginase (EC 3.5.1.1)
SAOUHSC_01812	-2,0	0,001	3'-to-5' oligoribonuclease A, Bacillus type (FIG146085)
SAOUHSC_01644	-2,0	0,001	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)
SAOUHSC_01501	-2,0	0,004	Elastin binding protein EbpS
SAOUHSC_01267	-2,0	0,005	2-oxoglutarate oxidoreductase, beta subunit (EC 1.2.7.3)
SAOUHSC_01807	-2,0	0,000	6-phosphofructokinase (EC 2.7.1.11)
SAOUHSC_00712	-2,0	0,002	oxidoreductase of aldo/keto reductase family, subgroup 1
SAOUHSC_00480	-2,0	0,001	Nucleoside triphosphate pyrophosphohydrolase MazG (EC 3.6.1.8)
SAOUHSC_01038	-2,0	0,002	Peptide deformylase (EC 3.5.1.88)
SAOUHSC_01259	-2,0	0,018	Hypothetical protein (FIG01108279)
SAOUHSC_02812	-2,0	0,001	Hypothetical protein (FIG01108416)
SAOUHSC_01190	-2,0	0,002	Thiamin pyrophosphokinase (EC 2.7.6.2)
SAOUHSC_00462	-2,0	0,003	Putative deoxyribonuclease YcfH
SAOUHSC_01815	-2,0	0,002	metal-dependent hydrolase (FIG146085)
SAOUHSC_01121	2,0	0,027	Alpha-hemolysin precursor
SAOUHSC_00135	2,0	0,002	FIG01108032: hypothetical protein
SAOUHSC_01919	2,0	0,023	Hypothetical protein (FIG001583), contains S4-like RNA binding domain
SAOUHSC_02243	2,0	0,005	Leukocidin LukS-PV
SAOUHSC_00608	2,0	0,006	Alcohol dehydrogenase (EC 1.1.1.1)
SAOUHSC_02629	2,0	0,051	Membrane component of multidrug resistance system
SAOUHSC_00535	2,0	0,017	L-threonine 3-dehydrogenase (EC 1.1.1.103)
SAOUHSC_01366	2,1	0,004	Anthranilate synthase, aminase component (EC 4.1.3.27)
SAOUHSC_00483	2,1	0,008	S1 RNA binding domain protein
SAOUHSC_02442	2,1	0,024	Hypothetical protein
SAOUHSC_02323	2,1	0,006	Cardiolipin synthetase (EC 2.7.8.-)
SAOUHSC_00258	2,1	0,008	Putative secretion accessory protein EsaA/YueB
SAOUHSC_02026	2,2	0,000	Hypothetical protein, phi-ETA orf58 homolog
SAOUHSC_00691	2,2	0,018	Undecaprenyl-diphosphatase (EC 3.6.1.27)
SAOUHSC_01455	2,2	0,003	Uncharacterized protein similar to YpbR
SAOUHSC_02756	2,2	0,016	Hypothetical protein (FIG01108728)
SAOUHSC_01342a	2,3	0,005	Large-conductance mechanosensitive channel
SAOUHSC_02580	2,3	0,042	N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28) / Endo-beta-N-acetylglucosaminidase
SAOUHSC_02611	2,4	0,008	Hypothetical protein (FIG01108370)
SAOUHSC_01809	2,4	0,005	Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
SAOUHSC_00808	2,4	0,001	Hypothetical SAV0808 homolog, near pathogenicity islands SaPI att-site
SAOUHSC_00481	2,4	0,018	Ribosome-associated heat shock protein (S4 paralog)

Continued in the following page

Annex I. Continued.

Gene ID	Fold change	Anova (p)	PATRIC Description
SAOUHSC_00660	2,4	0,016	Hypothetical protein (FIG01108275)
SAOUHSC_02463	2,5	0,054	Hyaluronate lyase precursor (EC 4.2.2.1)
SAOUHSC_02407	2,5	0,006	Hypothetical protein similar to YbbP, contains nucleotide-binding domain
SAOUHSC_00429	2,6	0,001	Hypothetical protein (FIG01107849)
SAOUHSC_02698	2,6	0,020	L-Cystine ABC transporter, permease protein TcyB
SAOUHSC_01002	2,6	0,006	quinol oxidase polypeptide II QoxA (EC:1.9.3.-)
SAOUHSC_01680	2,6	0,020	Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.-)
SAOUHSC_01796	2,6	0,022	Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
SAOUHSC_01110	2,7	0,003	Hypothetical protein, similarity with fibrinogen-binding protein Efb
SAOUHSC_02571	2,7	0,013	Secretory antigen precursor SsaA
SAOUHSC_00249	2,7	0,015	ABC transporter ATP-binding protein
SAOUHSC_01633	2,7	0,002	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P1 protein) (EC 1.4.4.2)
SAOUHSC_01338	2,7	0,008	alternate gene name: yoxG
SAOUHSC_01462	2,7	0,009	Cell division protein GpsB
SAOUHSC_01838	2,7	0,001	Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
SAOUHSC_00069	2,8	0,038	Protein A, von Willebrand factor binding protein Spa
SAOUHSC_02028	2,8	0,002	Hypothetical protein (FIG01108548)
SAOUHSC_02127	2,8	0,000	Staphopain A precursor (EC 3.4.22.48)
SAOUHSC_01241	2,9	0,004	DNA polymerase III polC-type (EC 2.7.7.7)
SAOUHSC_00264	2,9	0,001	EsaC protein within ESAT-6 gene cluster
SAOUHSC_00256	2,9	0,002	Secretory antigen precursor SsaA
SAOUHSC_00427	2,9	0,002	Autolysin precursor
SAOUHSC_02706	3,0	0,000	IgG-binding protein SBI
SAOUHSC_00728	3,0	0,000	Lipoteichoic acid synthase LtaS Type Ib
SAOUHSC_02839	3,1	0,007	L-serine dehydratase, alpha subunit (EC 4.3.1.17)
SAOUHSC_01183	3,2	0,000	Methionyl-tRNA formyltransferase (EC 2.1.2.9)
SAOUHSC_02816	3,3	0,014	Alkaline phosphatase like protein
SAOUHSC_00648	3,3	0,013	nucleoside transport protein
SAOUHSC_01418	3,3	0,027	2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)
SAOUHSC_01714	3,3	0,019	Transcription elongation factor GreA
SAOUHSC_01356	3,4	0,026	Transcription antiterminator
SAOUHSC_02029	3,5	0,042	Putative major teichoic acid biosynthesis protein C
SAOUHSC_02680	3,6	0,002	Respiratory nitrate reductase beta chain (EC 1.7.99.4)
SAOUHSC_01827	3,8	0,004	Septation ring formation regulator EzrA
SAOUHSC_00300	3,9	0,000	Triacylglycerol lipase (EC 3.1.1.3)
SAOUHSC_02541	4,0	0,000	Molybdopterin-guanine dinucleotide biosynthesis protein MobB
SAOUHSC_02099	4,0	0,002	Sensor histidine kinase VraS
SAOUHSC_02241	4,0	0,001	Leukocidin LukF-PV
SAOUHSC_02972	4,1	0,001	immunodominant antigen B
SAOUHSC_02971	4,1	0,003	Zinc metalloproteinase precursor (EC 3.4.24.29) / aureolysin
SAOUHSC_00421	4,2	0,001	Cystathionine beta-synthase (EC 4.2.1.22)
SAOUHSC_01114	4,2	0,002	Extracellular fibrinogen-binding protein Efb
SAOUHSC_00562	4,3	0,000	Hydroxymethylpyrimidine phosphate kinase ThiD (EC 2.7.4.7)
SAOUHSC_02855	4,3	0,028	Secretory antigen SsaA
SAOUHSC_01964	4,6	0,001	Uncharacterized protein, homolog of B.subtilis yhgC
SAOUHSC_02240	4,6	0,038	Beta-hemolysin
SAOUHSC_01039	4,7	0,003	Hypothetical protein (FIG01108153)
SAOUHSC_01759	5,2	0,006	Rod shape-determining protein MreC
SAOUHSC_00897	5,6	0,001	Glycerophosphoryl diester phosphodiesterase, periplasmic (EC 3.1.4.46)
SAOUHSC_00994	5,8	0,001	N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)/ endo-beta-N-acetylglucosaminidase
SAOUHSC_02883	6,2	0,012	Secretory antigen precursor SsaA
SAOUHSC_00803	6,3	0,001	3'-to-5' exoribonuclease RNase R
SAOUHSC_02064	6,4	0,002	Phage protein
SAOUHSC_03004	7,5	0,000	Polysaccharide intercellular adhesin (PIA) biosynthesis deacetylase IcaB
SAOUHSC_01108	7,8	0,002	phosphoesterase
SAOUHSC_00819	11,0	0,001	Cold shock protein CspC
SAOUHSC_00818	13,3	0,000	Phage-encoded chromosome degrading nuclease YokF
SAOUHSC_02710	218,0	0,000	Gamma-hemolysin component B

Annex II. CspA^{3xFLAG} RIP-chip results. The genomic coordinates (Summit_position) and the tiling intensity signals (Summit_value) for the CspA-binding peaks are indicated. Boundaries, length and features of the CspA-targeted transcripts as well as description of the closest feature to the corresponding peak is provided.

Summit position	Summit value	Peak strand	Transcript			Features in Transcript	Feature closer to peak	Gene Name	Description
			Start	End	Length				
12640	1084,8	+	12473	14185	1712	T_box serS	00009	serS	Seryl-tRNA synthetase (EC 6.1.1.11)
115267	650,8	-	115195	115606	411	ncRNA	ncRNA		non-coding RNA
140838	150,5	+	140777	141609	832	00135	00135	00135	FIG01108032: hypothetical protein
157659	1121,2	-	157136	157723	587	00146	00146	00146	FIG01107837: hypothetical protein, integral membrane protein
168271	121,5	-	166616	169004	2388	rsaK pstG	00155	ptsG	PTS system glucose-specific protein (EC 2.7.1.69)
180213	585,0	-	180178	182034	1856	00167	00167	00167	Putative glutathione transporter, ATP-binding component
228464	3037,0	+	228171	229390	1219	00206	00206	00206	L-lactate dehydrogenase (EC 1.1.1.27)
254742	145,0	+	254103	255353	1250	00232 00233	00233	00233	murein hydrolase export regulator, antiholin-like protein LrgB
275952	782,0	+	275857	276274	417	esxA	00257	esxA	ESAT-6/Esx family secreted protein EsxA/YukE
297001	298,0	-	296288	297197	909	00281	00281	00281	formate/nitrite transporter family protein
298485	245,0	-	297358	298802	1444	brnQ	00282	brnQ	Branched-chain amino acid transport system carrier protein
303028	178,8	+	302323	303501	1178	00290	00290	00290	Perfringolysin O regulator protein PfoR
341871	250,2	-	341658	343238	1580	00330 00329 00328	00328	00328	hypothetical protein, MttB family protein
345742	234,0	+	343403	345957	2554	00331 00332 00333 00334	00334	00334	conserved hypothetical protein
359994	171,0	+	357853	360175	2322	ychF	00346	ychF	GTP-binding and nucleic acid-binding protein YchF
360638	880,3	+	360547	361788	1241	rpsF 00349 rpsR	00348	rpsF	ribosomal protein S6
361184	265,2	+	360547	361788	1241	rpsF 00349 rpsR	00349	00349	Single-stranded DNA-binding protein
361492	289,0	+	360547	361788	1241	rpsF 00349 rpsR	00350	rpsR	ribosomal protein S18, zinc-independent
361947	818,8	-	361903	363339	1436	00352 00351 ncRNA	00351	00351	conserved hypothetical protein
366784	821,0	+	366659	366997	338	00358	00358	00358	transglycosylase-associated protein
410219	191,5	-	410048	410578	530	00409	00409	00409	conserved domain protein
421020	205,2	+	420781	422212	1431	00420	00420	00420	Sodium-dependent transporter
422504	215,0	+	422300	424497	2197	00421 00422	00421	00421	Cystathionine beta-synthase (EC 4.2.1.22)
424310	219,0	+	422300	424497	2197	00421 00422	00422	00422	Cystathionine gamma-lyase (EC 4.4.1.1)
428370	346,5	+	427450	428795	1345	00427	00427	00427	Autolysin precursor, N-acetylmuramoyl-L-alanine amidase
430491	223,5	-	430189	432177	1988	00433 00431	00431	00431	putative membrane protein
447886	231,8	+	446995	448042	1047	00444 recR	00445	recR	Recombination protein RecR
456790	408,5	+	456596	459427	2831	00452 00454 00455 00456	00452	00452	conserved hypothetical protein
466436	358,5	+	466181	466560	379	00465	00465	00465	Veg protein
468508	134,0	+	468334	469319	985	00468 00469	00468	00468	putative endoribonuclease L-PSP
472148	245,8	+	472089	472924	835	00474	00474	00474	LSU ribosomal protein L25p
487548	451,0	+	486656	487761	1105	00488	00488	cysK	Cysteine synthase (EC 2.5.1.47)
502087	303,0	-	501724	503037	1313	00501	00501	00501	Nucleoside permease NupC
512916	530,5	+	512707	512995	288	Ribosw cysS	00510 (Ribosw)	cysE	Serine acetyltransferase (EC 2.3.1.30)

Continued in the following page

Annex II. Continued.

Summit position	Summit value	Peak strand	Transcript			Features in Transcript	Feature closer to peak	Gene Name	Description
			Start	End	Length				
518782	508,0	+	518560	520029	1469	rplK rplA	00518	rplK	LSU ribosomal protein L11p (L12e)
519958	754,0	+	518560	520029	1469	rplK rplA	00519	rplA	LSU ribosomal protein L1p (L10Ae)
520252	324,0	+	520063	521268	1205	Ribosw rplJ rplL	00520	00520	LSU ribosomal protein L10p (P0)
521022	189,2	+	520063	521268	1205	Ribosw rplJ rplL	00521	rplL	LSU ribosomal protein L7/L12 (P1/P2)
534196	1185,5	+	529574	534583	5009	00526 rpsL rpsG fusA tuf	00530	tuf	Translation elongation factor Tu
538088	156,0	+	537555	538526	971	00533	00533	00533	chaperone protein HchA
564506	226,0	+	563992	565609	1617	00556	00556	00556	L-Proline/Glycine betaine transporter ProP
565857	2186,0	-	565565	565902	337	ncRNA	ncRNA	NA	non-coding RNA
569301	4094,0	-	569178	569598	420	00561	00561	00561	conserved domain protein
572808	581,5	+	572730	574290	1560	00569	00569	00569	cationic amino acid transporter
575902	291,5	+	575845	576120	275	rsaA	RsaA	RsaA	non-coding RNA
609908	185,0	+	609295	610142	847	00618	00618	00618	FIG01108193: hypothetical protein
623747	367,5	-	623345	624458	1113	RsaC	RsaC	RsaC	non-coding RNA
626127	570,0	-	623345	627121	3776	00637 00636 00635 rsaC	00636	00636	iron (chelated) ABC transporter permease
626981	544,5	-	623345	627121	3776	00637 00636 00635 rsaC	00637	00637	Manganese ABC transporter, ATP-binding protein SitB
631650	200,0	+	630622	631749	1127	00642	00642	00642	Teichoic acid translocation permease protein TagG
646630	518,3	+	646310	646921	611	00659	00659	00659	conserved hypothetical protein
647344	273,5	+	647276	648470	1194	00660	00660	00660	branched-chain amino acid transport system II carrier protein
658082	231,0	+	657120	659012	1892	00669 00670	00670	00670	phosphate transporter
659615	232,0	-	659463	660418	955	00671	00671	00671	conserved hypothetical protein
669716	312,0	+	666322	669952	3630	00678 00679 00680 00681 00682	00682	00682	conserved hypothetical protein
675169	171,0	-	674752	675735	983	00691	00691	00691	Undecaprenyl-diphosphatase (EC 3.6.1.27)
679789	288,0	-	679202	680075	873	00694	00694	mgrA	Transcriptional regulator MgrA
693390	777,0	+	690231	694053	3822	00706 00707 00708	00708	00708	fructose specific permease PTS system, fructose-specific (EC 2.7.1.69)
702651	245,3	-	702383	703056	673	00718	00718	00718	putative membrane protein
745372	412,0	+	744206	745564	1358	00762	00762	00762	probable undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphatetransferase
774338	9787,5	+	774294	774417	123	RsaH	RsaH	RsaH	non-coding RNA
778230	292,5	+	776018	778299	2281	00794 00795	00795	gap	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
784334	773,3	+	784200	784489	289	00801	00801	secG	preprotein translocase, SecG subunit
801204	850,2	+	800965	801321	356	00819	00819	cspC	Cold shock protein CspC
803311	214,0	-	803125	803693	568	00824 00823	00823	00823	conserved hypothetical protein
803983	684,5	-	803896	804470	574	00826 (Overlapping 3')	00826	00826	conserved hypothetical protein
816212	180,3	+	815991	816314	323	00845	00845	00845	conserved hypothetical protein
816821	277,5	-	816348	817371	1023	00846	00846	00846	integral membrane protein
838675	121,5	-	838365	838917	552	00873	00873	00873	conserved hypothetical protein, nitrogen-fixing NifU, C-terminal
842966	240,0	+	841625	843093	1468	00878	00878	00878	NADH dehydrogenase (EC 1.6.99.3)
876538	146,7	+	875640	876621	981	00906	00906	00906	Fumarylacetoacetate hydrolase family protein
880122	484,5	+	879738	880230	492	00910	00910	00910	N-6 Adenine-specific DNA methylase YitW

Continued in the following page

Annex II. Continued.

Summit position	Summit value	Peak strand	Transcript			Features in Transcript	Feature closer to peak	Gene Name	Description
			Start	End	Length				
890755	589,5	-	890591	890881	290	00919	00919	00919	conserved hypothetical protein
893282	309,5	+	891111	893391	2280	00920 00921	00921	00921	3-oxoacyl-[acyl-carrier-protein] synthase, KASII (EC 2.3.1.41)
907170	229,0	+	906625	907303	678	00934	00934	spxA	regulatory protein spxA
908304	303,0	+	907421	908372	951	00935	00935	00935	Negative regulator of genetic competence MecA
920722	184,0	+	919955	920853	898	00947	00947	00947	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)
925370	139,0	+	925061	925674	613	00951	00951	00951	2H phosphoesterase superfamily protein Bsu1186 (yjcG)
932706	199,0	+	931872	932923	1051	yybP ykoY 00957	00957	00957	conserved hypothetical protein, toxic anion resistance protein, integral membrane protein TerC
941204	511,5	+	941116	942032	916	00964 00965	00964	00964	FIG01108367: hypothetical protein
942170	178,3	+	941116	942032	916	00964 00965	00965	00965	hypothetical protein
949051	249,2	-	948687	949167	480	00975	00975	00975	hypothetical protein
952005	182,0	-	951760	952777	1017	00980	00980	menA	1,4-dihydroxy-2-naphthoate octaprenyltransferase (EC 2.5.1.74)
957192	151,5	+	956703	957611	908	00985	00985	menB	Naphthoate synthase (EC 4.1.3.36)
972599	786,0	-	972359	976602	4243	01002 01001 01000 00999	00999	qoxD	cytochrome aa3 quinol oxidase, subunit IV
973369	689,5	-	972359	976602	4243	01002 01001 01000 00999	01000	qoxC	cytochrome aa3 quinol oxidase, subunit III
974125	295,5	-	972359	976602	4243	01002 01001 01000 00999	01001	qoxB	quinol oxidase polypeptide I QoxB (EC:1.9.3.-)
976589	157,7	-	972359	976602	4243	01002 01001 01000 00999	01002	01002	quinol oxidase polypeptide II QoxA (EC:1.9.3.-)
977387	316,5	-	977080	977534	454	01005	01005	01005	chitinase-related protein
995762	624,0	+	994363	995892	1529	01025	01025	01025	FIG01107839: hypothetical protein
999346	198,0	+	997401	1000188	2787	01027 01028 01029	01029	ptsI	Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)
1033401	202,5	-	1032357	1033672	1315	01065	01065	01065	heme A synthase
1034234	203,0	+	1033726	1035261	1535	01066 01067	01066	cyoE	protoheme IX farnesyltransferase
1034864	591,3	+	1033726	1035261	1535	01066 01067	01067	01067	hypothetical protein
1041276	296,3	+	1041239	1042135	896	01077 01078	01077	01077	conserved hypothetical protein
1042046	1897,7	+	1041239	1042135	896	01077 01078	01078	rpmF	ribosomal protein L32
1062024	295,5	+	1061895	1062299	404	01100	01100	trx	Thioredoxin
1093874	236,0	+	1092078	1093963	1885	01142 01143 01144	01144	01144	Cell division protein FtsL
1106992	299,8	+	1106948	1109980	3032	T-box 01159	01159_T-box	ileS	Isoleucyl-tRNA synthetase (EC 6.1.1.5)
1114006	292,0	+	1113379	1123855	10476	01164 01165 pyrB pyrA 01168 01169 carB pyrF pyrE 01173	01164	01164	Uracil phosphoribosyltransferase (EC 2.4.2.9) / Pyrimidine operon regulatory protein PyrR
1127628	178,0	+	1126795	1127787	992	01176 01177	01177	rpoZ	DNA-directed RNA polymerase, omega subunit

Continued in the following page

Annex II. Continued.

Summit position	Summit value	Peak strand	Transcript			Features in Transcript	Feature closer to peak	Gene Name	Description
			Start	End	Length				
1133193	460,5	-	1133021	1133386	365	01181	01181	01181	hypothetical protein
1143133	1187,0	-	1142888	1143211	323	01191	01191	rpmB	ribosomal protein L28
1151008	170,0	+	1148068	1151823	3755	01196 01197 01198 01199 01201	01199	fabG	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
1151624	674,0	+	1151508	1151823	315	01201	01201	acpP	Acyl carrier protein / HmrB protein involved in methicillin resistance
1161522	1499,5	+	1161457	1161988	531	01211	01211	rplS	LSU ribosomal protein L19p
1185672	309,0	+	1184400	1186087	1687	01237 01238	01238	cdsA	Phosphatidate cytidylyltransferase (EC 2.7.7.41)
1201408	457,0	+	1201340	1201774	434	01250	01250	rpsO	SSU ribosomal protein S15p (S13e)
1219132	377,0	+	1217544	1219217	1673	01263	01263	01263	FIG002344: Hydrolase (HAD superfamily)
1269469	288,0	-	1268315	1269849	1534	01326	01326	01326	Lysine-specific permease
1271618	371,0	+	1269992	1271856	1864	01327 01328	01328	rpmG	ribosomal protein L33
1278814	177,2	+	1278729	1279112	383	01338	01338	01338	alternate gene name: yoxG
1279346	253,8	+	1279199	1279766	567	01340	01340	01340	hypothetical protein
1285562	334,0	+	1284536	1286345	1809	01346	01346	01346	Glycine betaine transporter OpuD
1291183	311,0	-	1290679	1291296	617	01349	01349	01349	FIG01108262: hypothetical protein
1291953	410,0	-	1291334	1292026	692	01350	01350	01350	Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY
1300864	537,0	+	1298789	1301066	2277	01356 01357 01358	01358	01358	hypothetical protein
1310048	233,0	+	1309993	1316880	6887	T_box05 T_box06 01366 01367 01368 01369 01370 01371 01372	01366_T-box	01366	Anthranilate synthase, aminase component (EC 4.1.3.27)
1345965	2415,0	-	1345800	1346231	431	01403	01403	cspA	Cold shock protein CspA
1351005	255,8	-	1349764	1351237	1473	01411	01411	brnQ	branched-chain amino acid transport system II carrier protein
1375176	258,5	+	1374326	1375556	1230	01440	01440	01440	hypothetical protein
1423294	141,0	+	1420791	1423796	3005	01466 01467	01467	01467	Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-) / Penicillin-binding protein 1A/1B (PBP1)
1434991	1487,2	-	1434764	1435559	795	01477	01477	01477	probable Zn-dependent protease
1445659	2315,0	-	1445568	1445924	356	01490	01490	01490	DNA-binding protein HU-beta
1458441	177,0	-	1457987	1458849	862	RFN 01505	01505	01505	riboflavin transporter
1513489	243,0	-	1511665	1514226	2561	01586 01585	01586	SrrA	DNA-binding response regulator SrrA
1551478	158,0	+	1551210	1551873	663	01630	01630	01630	Rhodanese-like domain protein
1556007	962,0	-	1551836	1555930	4094	gcvT_Ribosw 01634 01633 01632	01634_gcvT	gcvT	Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)
1567319	292,0	-	1567094	1567840	746	01653	01653	01653	Manganese superoxide dismutase (EC 1.15.1.1); Superoxide dismutase [Fe] (EC 1.15.1.1)
1579352	240,8	+	1579250	1580952	1702	RsaOD 01666	01666_RsaOD	glyS	Glycyl-tRNA synthetase (EC 6.1.1.14)
1587563	313,3	-	1585128	1587617	2489	01677 01676 01675	01677	01677	hypothetical protein
1587871	382,5	-	1587726	1588060	334	01678	01678	rpsU	SSU ribosomal protein S21p
1600254	124,0	+	1599925	1600400	475	01689	01689	rpsT	SSU ribosomal protein S20p
1624201	233,8	-	1624100	1625228	1128	01721 01720 01719	01719	01719	hypothetical protein

Continued in the following page

Annex II. Continued.

Summit position	Summit value	Peak strand	Transcript			Features in Transcript	Feature closer to peak	Gene Name	Description
			Start	End	Length				
1628121	328,5	-	1624100	1628201	4101	T_box07 01722 01721 01720 01719	01722_T box	01722	Alanyl-tRNA synthetase (EC 6.1.1.7)
1638061	227,0	-	1637960	1638876	916	01735	01735	01735	HesA/MoeB/ThiF family protein related to EC-YgdL
1642639	252,5	-	1639007	1642772	3765	Ribosw 01738 01737	01738	hisS	Histidyl-tRNA synthetase (EC 6.1.1.21)
1639181	4957,5	-	1639007	1639237	230	ncRNA	ncRNA	NA	non-coding RNA
1652607	312,8	-	1652284	1653522	1238	01748 01747	01747	yajC	Preprotein translocase subunit YajC (TC 3.A.5.1.1)
1658613	1211,0	-	1658526	1659659	1133	01757 01756 01755	01755	rpmA	LSU ribosomal protein L27p
1659551	182,0	-	1658526	1659659	1133	01757 01756 01755	01757	rplU	LSU ribosomal protein L21p
1662232	660,8	+	1662033	1662338	305	01761a	01761a	01762a	hypothetical protein
1662449	283,0	-	1662293	1662691	398	01762	01762	01762	conserved hypothetical protein
1668581	470,5	-	1664327	1668656	4329	Ribosw 01767 01766	01767_Ri bosw	valS	Valyl-tRNA synthetase (EC 6.1.1.9)
1670807	139,5	-	1670716	1676942	6226	01776 01775 01774 01773 01772 01771	01771	hemL	Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
1682567	231,0	-	1682409	1683776	1367	01786 01785 01784	01784	rplT	LSU ribosomal protein L20p
1687915	330,0	-	1685424	1688075	2651	T_box08 01788	01788_T-box	thrS	Threonyl-tRNA synthetase (EC 6.1.1.3)
1705646	457,5	+	1704258	1705724	1466	01803	01803	01803	amino acid permease-associated region
1708551	154,5	-	1707914	1710747	2833	01807 01806	01806	pyk	Pyruvate kinase (EC 2.7.1.40)
1721809	127,3	-	1721229	1722034	805	01815	01815	01815	FIG002379: metal-dependent hydrolase
1734430	345,0	+	1733840	1735221	1381	01828 01829	01829	rpsD	SSU ribosomal protein S4p (S9e)
1746155	481,0	-	1744592	1746232	1640	T_box09 01839	01839_T-box	tyrS	Tyrosyl-tRNA synthetase (EC 6.1.1.1)
1773427	156,3	-	1771835	1773590	1755	Ribosw 01866 01865	01866_Ri bosw	01866	amino acid permease-associated region
1776311	616,0	-	1776218	1777536	1318	01869	01869	01869	conserved hypothetical protein
1791389	378,5	-	1788410	1791460	3050	T_box10 01875 01874	01875_T-box	leuS	Leucyl-tRNA synthetase (EC 6.1.1.4)
1806901	203,0	-	1806697	1807686	989	01895	01895	01895	Autolysin (EC 3.5.1.28)
1857847	256,5	-	1857709	1857931	222	01953	01953	01953	conserved domain protein
1870496	257,5	+	1870084	1870736	652	01964	01964	01964	Signal transduction protein TRAP
1873744	218,5	+	1873288	1873872	584	01969	01969	01969	conserved hypothetical protein
1971100	210,0	+	1970578	1971509	931	02095 02096	02096	02096	FIG010063: hypothetical protein
2018791	233,5	-	2018612	2019226	614	02145	02145	02145	FIG01108203: hypothetical protein
2027009	438,5	-	2026887	2027175	288	02156	02156	02156	conserved domain protein
2131967	242,5	-	2131841	2133553	1712	02300 02299 02298	02298	sigB	RNA polymerase sigma factor SigB
2154052	228,0	+	2152935	2154323	1388	02319	02319	02319	cell cycle protein
2154549	450,0	-	2154256	2154595	339	02320	02320	02320	conserved hypothetical protein
2158385	419,3	-	2158271	2159248	977	02327	02327	02327	Inner membrane protein translocase component YidC, short form OxaI-like

Continued in the following page

Annex II. Continued.

Summit position	Summit value	Peak strand	Transcript			Features in Transcript	Feature closer to peak	Gene Name	Description
			Start	End	Length				
2167471	256,0	-	2167394	2178187	10793	02355 02354 02353 02352 02351 02350 02349 02347 02346 02345 02343 02341 02340	02340	atpC	ATP synthase epsilon chain (EC 3.6.3.14)
2182857	732,0	-	2182547	2182920	373	02361	02361	rpmE	LSU ribosomal protein L31p @ LSU ribosomal protein L31p, zinc-independent
2188779	475,0	-	2188266	2189248	982	02366	02366	fba	Fructose-bisphosphate aldolase class II (EC 4.1.2.13)
2190123	304,5	-	2190001	2193569	3568	02370 02369 02368	02368	pyrG	CTP synthase (EC 6.3.4.2)
2194274	167,0	+	2193898	2194786	888	02371	02371	coaA	Pantothenate kinase type II, eukaryotic (EC 2.7.1.33)
2199251	268,3	-	2199172	2202081	2909	02379 02378 02377 02376	02376	02376	conserved hypothetical protein
2206601	129,5	-	2206422	2207872	1450	02386 02385 02384	02384	02384	conserved domain protein
2250197	269,0	-	2249857	2252359	2502	02424 02424 02422	02422	02422	conserved domain protein
2271239	136,0	-	2270541	2272915	2374	02448 02447	02447	02447	Putative oxidoreductase YncB
2299379	492,0	-	2299275	2300374	1099	02478 02477	02477	rpsI	SSU ribosomal protein S9p (S16e)
2304125	159,5	-	2304011	2318343	14332	rpsJ, rplC, rplD, rplW, rplB, rplV rpsC, rplP, rpmC, rpsQ, rplN, rplX, rplE, rpsN, rpsH, rplF, rplR, rpsE, rpmD, rplO, secY, adk, infA, rpmJ, rpsM, rpsK, rpoA,rplQ	02484	rplQ	LSU ribosomal protein L17p
2315381	199,5	-	2304011	2318343	14332	Same than previous	02509	rplB	LSU ribosomal protein L2p (L8e)
2318636	346,2	+	2318477	2319077	600	02515	02515	02515	conserved hypothetical protein
2320071	553,0	-	2319029	2320438	1409	02516	02516	02516	Hypoxanthine/guanine permease PbuG
2324096	214,5	+	2323860	2325155	1295	02520	02520	02520	glucose uptake protein GlcU
2326588	174,5	+	2326200	2327141	941	02523 02524	02524	02524	conserved hypothetical protein
2362540	742,5	+	2361899	2362857	958	02571	02571	02571	Secretory antigen precursor SsaA2
2366964	311,5	+	2366634	2367315	681	02576	02576	02576	N-acetylmuramoyl-L-alanine amidase aaa
2378955	1178,5	-	2378886	2379728	842	02587	02587	02587	conserved hypothetical protein, abortive infection protein
2383085	214,0	-	2381516	2383267	1751	02590	02590	02590	Histidine transport protein (permease)
2391597	527,0	-	2389457	2391678	2221	S_box04 02601 02600	02601_S-box	02601	Na+/H+ antiporter family protein
2416454	117,3	+	2415846	2416787	941	02628	02628	02628	FIG01108339: hypothetical protein
2433177	322,5	-	2432948	2434729	1781	02648	02648	02648	L-lactate permease
2441241	384,5	-	2441078	2441621	543	02656	02656	02656	conserved hypothetical protein
2452077	259,2	-	2451990	2453353	1363	02667	02667	02667	proton/sodium-glutamate symport protein
2454702	313,7	+	2454255	2455435	1180	02669 02670	02669	sarZ	Transcriptional regulator SarZ (Staphylococcal accessory regulator Z)

Continued in the following page

Annex II. Continued.

Summit position	Summit value	Peak strand	Transcript			Features in Transcript	Feature closer to peak	Gene Name	Description
			Start	End	Length				
2455787	160,3	-	2455390	2456770	1380	02671	02671	02671	nitrite extrusion protein
2473595	328,0	-	2473309	2474239	930	02687	02687	02687	formate/nitrite transporter family protein
2480623	219,0	-	2480484	2482798	2314	02699 02698 02697	02697	02697	Molybdenum transport ATP-binding protein ModC (TC 3.A.1.8.1)
2506292	250,5	+	2505521	2506818	1297	02725	02725	02725	Bicyclomycin resistance protein TcaB
2519207	292,0	-	2517734	2520481	2747	02740 02739	02740	02740	putative transporter
2555978	152,0	+	2555360	2556400	1040	02779	02779	02779	FIG01107931: hypothetical protein
2555831	8117,0	-	2555765	2556195	430	02781	02781	02781	conserved hypothetical protein
2595500	326,5	+	2595214	2596028	814	02816	02816	02816	Alkaline phosphatase like protein
2627518	1424,2	+	2627373	2627682	309	ncRNA	ncRNA	NA	non-coding RNA
2628162	109,5	+	2627974	2628492	518	02855	02855	02855	Secretory antigen SsaA
2633349	200,5	-	2632758	2633443	685	02861	02861	ogt	methylated-DNA--protein-cysteine methyltransferase
2661335	499,0	-	2660606	2661491	885	02887	02887	isaA	Immunodominant staphylococcal antigen A precursor
2679416	169,0	+	2679205	2679765	560	02912	02912	02912	PhnB protein; putative DNA binding 3-demethylubiquinone-9 3-methyltransferase domain protein
2721199	858,5	-	2721111	2721871	760	02961	02961	02961	Transcriptional regulator, MarR family
2738986	213,0	+	2738054	2741067	3013	02975 02976	02976	manA	Mannose-6-phosphate isomerase (EC 5.3.1.8)
2774161	417,0	-	2774060	2775082	1022	03001	03001	icaR	biofilm operon icaADBC negative transcriptional regulator, IcaR
2778522	267,5	+	2775142	2778667	3525	03002 03003 03004 03005	03005	icaC	Polysaccharide intercellular adhesin (PIA) biosynthesis protein IcaC
2805976	194,3	+	2804232	2806165	1933	03032 03033	03033	03033	high-affinity nickel-transport protein NixA
2807201	352,5	-	2807132	2807764	632	03035	03035	03035	putative membrane protein
2813074	1024,0	+	2812935	2813326	391	03045	03045	cspB	Cold shock protein CspB
2821047	1635,5	-	2820979	2821291	312	03055	03055	rpmH	ribosomal protein L34

Annex III: Table summarizing the list of putative mRNAs targeted by CspA for which the protein levels were found significantly affected in the Δ*cspA* mutant strain.

Start	End	Length	Strand	Summit pos.	Gene ID	Fold change	Anova (p)	PATRIC Description
140777	141609	832	+	140838	SAOUHSC_00135	2,0	0,002	FIG01108032: hypothetical protein
422300	424497	2197	+	422504	SAOUHSC_00421	4,2	0,001	Cystathionine beta-synthase (EC 4.2.1.22)
427450	428795	1345	+	428370	SAOUHSC_00427	2,9	0,002	Autolysin precursor
456596	459427	2831	+	456790	SAOUHSC_00454	-2,6	0,024	DNA polymerase III delta prime subunit
468334	469319	985	+	468508	SAOUHSC_00469	-2,4	0,000	Putative role in sporulation (SpoVG)
520063	521268	1205	+	520252	SAOUHSC_00520	-2,4	0,010	LSU ribosomal protein L10p (P0)
537555	538526	971	+	538088	SAOUHSC_00533	-2,3	0,000	chaperone protein HchA
563992	565609	1617	+	564506	SAOUHSC_00556	-5,4	0,034	L-Proline/Glycine betaine transporter ProP
647276	648470	1194	+	647344	SAOUHSC_00660	2,4	0,016	Hypothetical protein (FIG01108275)
674752	675735	983	-	675169	SAOUHSC_00691	2,2	0,018	Undecaprenyl-diphosphatase (EC 3.6.1.27)
776018	778299	2281	+	778230	SAOUHSC_00794	-2,5	0,002	Central glycolytic genes regulator
800965	801321	356	+	801204	SAOUHSC_00819	11,0	0,001	Cold shock protein CspC
875640	876621	981	+	876538	SAOUHSC_00906	-2,4	0,006	Fumarylacetoacetate hydrolase family protein
891111	893391	2280	+	893282	SAOUHSC_00920	-2,3	0,001	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII
925061	925674	613	+	925370	SAOUHSC_00951	-2,2	0,001	2H phosphoesterase superfamily protein
972359	976602	4243	-	976589	SAOUHSC_01002	2,6	0,006	quinol oxidase polypeptide II QoxA (EC:1.9.3.-)
1113379	1123855	10476	+	1114006	SAOUHSC_01166	-2,6	0,008	Aspartate carbamoyltransferase (EC 2.1.3.2)
1113379	1123855	10476	+	1114006	SAOUHSC_01172	-2,1	0,007	Orotate phosphoribosyltransferase (EC 2.4.2.10)
1113379	1123855	10476	+	1114006	SAOUHSC_01168	-2,0	0,021	Dihydroorotase (EC 3.5.2.3)
1148068	1151823	3755	+	1151008	SAOUHSC_01198	-3,2	0,000	Malonyl CoA-acyl carrier protein transacylase
1148068	1151823	3755	+	1151008	SAOUHSC_01199	-2,9	0,000	3-oxoacyl-[acyl-carrier protein] reductase
1151508	1151823	315	+	1151624	SAOUHSC_01201	-2,2	0,002	Acyl carrier protein
1201340	1201774	434	+	1201408	SAOUHSC_01250	-3,1	0,010	SSU ribosomal protein S15p (S13e)
1278729	1279112	383	+	1278814	SAOUHSC_01338	2,7	0,008	alternate gene name: yoxG
1298789	1301066	2277	+	1300864	SAOUHSC_01356	3,4	0,026	Transcription antiterminator
1309993	1316880	6887	+	1310048	SAOUHSC_01366	2,1	0,004	Anthranilate synthase, aminase component
1345800	1346231	431	-	1345965	SAOUHSC_01403	-43,9	0,000	Cold shock protein CspA
1551836	1555930	4094	-	1556007	SAOUHSC_01632	-2,1	0,000	Glycine cleavage system P-protein
1551836	1555930	4094	-	1556007	SAOUHSC_01634	-2,1	0,014	Aminomethyltransferase
1551836	1555930	4094	-	1556007	SAOUHSC_01633	2,7	0,002	Glycine cleavage system P-protein
1567094	1567840	746	-	1567319	SAOUHSC_01653	-2,0	0,001	Manganese superoxide dismutase (EC 1.15.1.1);
1670716	1676942	6226	-	1670807	SAOUHSC_01773	-2,6	0,010	Uroporphyrinogen-III synthase (EC 4.2.1.75)
1670716	1676942	6226	-	1670807	SAOUHSC_01771	-2,3	0,002	Glutamate-1-semialdehyde aminotransferase
1682409	1683776	1367	-	1682567	SAOUHSC_01786	-2,9	0,004	Translation initiation factor 3
1707914	1710747	2833	-	1708551	SAOUHSC_01807	-2,0	0,000	6-phosphofructokinase (EC 2.7.1.11)
1721229	1722034	805	-	1721809	SAOUHSC_01815	-2,0	0,002	metal-dependent hydrolase (FIG146085)
1744592	1746232	1640	-	1746155	SAOUHSC_01839	-2,0	0,005	Tyrosyl-tRNA synthetase (EC 6.1.1.1)
1776218	1777536	1318	-	1776311	SAOUHSC_01869	-2,9	0,000	Hypothetical protein
1870084	1870736	652	+	1870496	SAOUHSC_01964	4,6	0,001	Homolog of *B. subtilis* yhgC
1873288	1873872	584	+	1873744	SAOUHSC_01969	-2,4	0,007	Hypothetical protein, S4-like RNA binding domain
2131841	2133553	1712	-	2131967	SAOUHSC_02298	-3,1	0,000	RNA polymerase sigma factor SigB
2131841	2133553	1712	-	2131967	SAOUHSC_02299	-2,2	0,000	Serine-protein kinase RsbW (EC 2.7.11.1)
2188266	2189248	982	-	2188779	SAOUHSC_02366	-3,5	0,000	Fructose-bisphosphate aldolase class II
2199172	2202081	2909	-	2199251	SAOUHSC_02377	-2,1	0,008	Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)
2270541	2272915	2374	-	2271239	SAOUHSC_02447	-2,3	0,002	Putative oxidoreductase YncB
2361899	2362857	958	+	2362540	SAOUHSC_02571	2,7	0,013	Secretory antigen precursor SsaA
2480484	2482798	2314	-	2480623	SAOUHSC_02698	2,6	0,020	L-Cystine ABC transporter, permease protein TcyB
2595214	2596028	814	+	2595500	SAOUHSC_02816	3,3	0,014	Alkaline phosphatase like protein
2627974	2628492	518	+	2628162	SAOUHSC_02855	4,3	0,028	Secretory antigen SsaA
2679205	2679765	560	+	2679416	SAOUHSC_02912	-2,1	0,004	PhnB protein
2721111	2721871	760	-	2721199	SAOUHSC_02961	-3,0	0,017	Transcriptional regulator, MarR family
2775142	2778667	3525	+	2778522	SAOUHSC_03004	7,5	0,000	Polysaccharide intercellular adhesin (icaB)